

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGGCAGTGGGCGGCACAGAGCAGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCTGTCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTATAGGACCGCCTAC
CGCCGCAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCT
CTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGA¹CTGCCCAGCGCCCCAGG
CTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCTCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAAATGAAA
CGTGAAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTL CVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTGCCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCCACGGCGGGGACTATGGTGAAATTCCC GGCGCTCACGCACTACTGGCCCCCTGATC
CGGTTCTTGGTGCCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGGCTGGCGT
ACTCCCTCATGAAGTTCTTACAGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCCTGCTCATCCCGATCCTCTCCTTGACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAGTGGCCCGGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCCAGCAACAACTGGTGAGCAGGCAACACAGTCACGGCAGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGCCCCCAGCTCTGTGCTGCG
GATCATCGTCTCATCGCCAGCCTCGTGGTCTACCTACCTGGGGGTGCACGGTGCAGCCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTGACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCCGTTCACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACCTGACATGGTTGAATGATAGCATTCTACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTGAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTCACAAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTGCGATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACCTGAAGCTTTAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACACCATAACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACATATGTGGTCTCCTGAAATAACGCATTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTTGAAAGCAGGTTTTTTCTTTAAAAAAATATAGACACGGTTCCT
AAATTGATTTAGTCAGAATTCCTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVFN SKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYII INKLHHVDES
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCLGYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQORISRPIVNL FVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSN TVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVvPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLG VHGATLGVG SLL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPPTEEVTDIVEMRENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCCTTTGCGGATTTTCT
CCTTTTTCCCAGTTCAGTCACAGNGAGGGCGCATCTCACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCGGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

Footnote: 48021001

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGG
CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG
AATGGGTTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCTGTTGGCTCCTTCTCTTCA
TCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGGCCGAGGAGTGCGATTCCCCTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT
CTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTCTGCCCAAGGTCCAGGACGCCCAGCCAACTCGGGTCTGCTGCAGGCCTCGGTCTAT
CACCTCTACACCATGTTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCTCTCTGTGCACCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT
GGCCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCTTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCTTCCCCTCCTCCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAACCTCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT
CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTTCTAATAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SOLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW
NQRWLGKAEECDRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQA
ACEGRAFDNEQDGVTYSSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICAS
WAGLLLLYLWTLVAPLLLLLRNDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG
TGA CTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGTATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATAACCGAAATAAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTACGGGAAGTCTTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAACAAGACAAGGACTTGAAAGTGCTCTGAACTTGAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTTCTTTTCTT
TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQOYIEQSQAIEYHNRFDAVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEFWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTTCGTTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 14

GAGCCGCCGCCGCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCGGCCCGCCCCACCCACGTCT
GCGTTGCTGCCCCGCCTGGGCCAGGCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTGCGGGG
GCTTCGCGTGTTC AAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTTAACC AATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCGCGCAATCCTAGTGCAATCCTTTGATGAGAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTTCTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCAATGTGCGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTTTTACCTGCAGAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC
TGGA AAAAGAGTGGAATTTATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTC CCAATTTTTTTTTTGGTCTTTT TAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT
CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTT CATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIIVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEXAGEVLRVVGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

404200F-4302400F

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCGGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

AATCCCAAATTCCCCAATTTTTTTGGNCTTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCTTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTC CGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 18

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCCTTCCACCTGATTGTGTCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTTCCAGCGCCAATTCTCAGAGCTGTACCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTGAGACAAAGTTGCCCTGCAGAGGTGAGCTGCCCCG
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCTCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCTCCTGTCCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCCAACAACGAAAAATAATCTGGCCCAAATGTGAGTTGTAAGTGTGAGTTTGAAGAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTCAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTCGCCAAACACCG
ACTCTGTCTGTTGCCTTGATCTTGAACCTCCAGCCTCCAGAAGTATGAGAAATAAAATTTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGV LWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAPKPVFEGDLLVLRCAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSSAAPPTLNPAPOKSAAPGTAPEEAPGPLPPPPTPSSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCTTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTTTGACCCGT
CCCTGCCCTCAATTTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLLTVDTYGRPILEVPESVTGPWKGDVNLPCITYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSYHTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRLSLQCQARGSPPISYIWKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFFVVKDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTDDMDGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

Downloaded from www.nrcresearchpress.com by University of New Brunswick on 09/11/15

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGGACGGGGCGCGGGGCGGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCGCGCCAGCTCGCCCGAGGTCCGTCCGA
GGCGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGGATGTCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCTGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
CCAGTGACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCTCCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCTCCACTCGCTCCACAGCAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACACACCCAGCATGATCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCTGAATTACAATGGACTTGACTCCACGCTTTCTAGGAGTCAGGGTCTTGGACTC
TTCTCGTCATTGGAGCTCAAGTACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTGAGATGAGCATTTTCTTATACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTGATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
TTCAAGAGGAAATGGGATGCTGTTGTAAATTTCTATGCATTTCTGCAAACCTATTGGATT
ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCACTCTATACTAATCTCACTATTTGTATTGAGCCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCTCAAAT
CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCAAT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTGLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTSYSGLYQCTAGNEAGKESCVVRVTQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 24

FIGURE 25

GTCTGTTCTTTGCTCTCTCGCGCCAGTCTCTCTCCCTGGTTCTCTCTCAGCCGCTGTCTGGAGGAGAGCACCCGGA
GACGCGGGCTGCAGTCGCGGCGGCTTCTCCCCGCTGGGCGGCTCGCCGCTGGGCGAGTGTCTGAGCGCCCTAG
AGCCTCCCTTGCCGCTCTCTCTCTGCCCCGCGCAGCAGTGCACATGGGGTGTGGAGGTAGATGGGCTCCCG
GCCCGGAGGCGGCGGTGGATGCGGCGCTGGGCGAAGCAGCCGCGGATTCCAGCTGCCCCGCGCGCCCCGGGCG
CCCCGCGAGTCCCCGGTTCAAGCATGGGGACCTCTCCGAGCAGCAGCACCGCCCTCGCCTCTGACCGGCATC
GCCCCGAGCCACAGCCACGATGATCGCGGGCTCCCTTCTCTCTGCTTGGATTCTTAGCACCACACAGCTCAG
CCAGAACAGAAGGCCTCGAATCTCATTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT
GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGACGAGTGTGCCCT
GTGGGACCTTTTACCAGGCATGAGAAATGCGATAGAGAAATGCCATGACTGTAGTCAGCCATGCCATGGCCAATG
ATTGAGAAATTACCTTGTGCTGCCCTGACTGACCGAGAATGCACCTTGCCACCTGGCATGTTCCAGTCTAACGCT
ACCTGTGCCCCCATAACGGTGTGTCTGTGGGTTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
TGTAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCCTTCTAGTGTGATGAAATGCCAAAGCATACACAGACTGT
CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCTCTTC
TCCAGTCCACCTCACCTTCCCTGGCACAGCCATCTTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT
TCCTCCACTTATGTTCCCAAAGGCATGAACCTCAACAGAATCCAACCTTCTGCTCTGTTAGACCAAAGGTACTG
AGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCCTC
CCAAACCTTCAGGTAGTCAACCACCAAGGCCCCCAACACAGACACATCCTGAAGCTGCTGCCGTCCATGGAG
GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG
CATTTTGACATCAATGAGCATTGTGCCCTGGATGATTGTGCTTTTCTGCTGCTGGTGTCTGTGGTGATTGTGGTG
TGCACTATCCGGAAGGCTCGAGGACTCTGAAAAAGGGGCCCCCGGAGGATCCAGTGCATTGTGAAAAAGGCA
GGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
ATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATTGCCAGTGAG
AGGAGGTTGTGCTTTTCTCAATGGGTACACAGCCGACCCAGCGGGCCTACGCAGCTCTGCGAGCATGGACC
ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAG
AAGATTGCTGGGCTGATGGAAGACACCACCAGCTGGAAACTGACAACTAGCTCTCCCGATGAGCCCGAGCCCG
CTTAGCCCGAGCCCCATCCCCAGCCCCAACCGGAAACTTGAGAATTCGCTCTCTCTGACGGTGGAGCCTTCCCCA
CAGGACAAGAACAAAGGGCTTCTTCGTGGATGAGTCGGAGCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC
TCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAGGACACAGTGTGCGGCAGGTACGCTGGAC
CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGAG
ATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTGAAATTTATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
CTCCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAATTAATCTCA
ATTAGTGGCAGGGTGGTTTTTAAATTTCTTCTGTTTCTGATTTTGTGTTTGGGGTGTGTGTGTGTGTGTGTGT
GT
TCTCTCTCTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTTATAAGCCTTTGCCAGGTGTAACCTGTGTGAA
ATACCCACCACTAAAGTTTTTTAAGTTCATATTTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTTCTG
TGCACTTTAAATTTACTTAACCTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC
TTCTTAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTATGTCTCTTAACATTACACCTACTTTTT
AAAAACAAATATTATTACTATTTTATTTATTTGTTTGTCTTTATAAATTTTCTTAAAGATTAAGAAAATTTAAGA
CCCCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTTCAATTTATGG
CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTTACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC
TCTTCTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT
GACAACTGGGCCACCAAGAACCTGAACTTCACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACTTT
GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTCCATAGAGAAATTTGCCAGCTTTGCTTTAAAGATGTCTTG
TTTTTTATATACATAATCAATAGGTCCAATCTGCTCTCAAGGCCCTGGTCTGGTGGGATTCCCTTACCAAT
ACTTTAATTAAAAATGGCTGCAACTGTAAGAACCTTGTCTGATATATTGCAACTATGCTCCCATTTACAAATG
TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTGTGG
GTAGTGGTGAAGGACCGATATCAGAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTA
AAAAAAA

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCC
TTGATGATTTGAAAATAAGTCTGTTCACACAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTACACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGAGTTT
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCACTGCACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCAGGGCTACCACCTGTGCGGGGGCTCTGTATCAC
GCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCTGTCTGAACCACGCGGCCGTCCCTTTGATTTCACAAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGGCCCTGTTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTACCTCCTTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAAC**TGAA**
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCT
GGACTCCCCTGTAGGAACCTGCACACGAGCAGACACCCTTGAGCTCTGAGTTCGGGCACCA
GTAGCAGGCCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAACTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGGAGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTTGTATTTTGTAGTAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTCCCCTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCTTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAAGTGCAGAACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCAAACCCACCCTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTCTT
TGTAGCATTGTTGGTGTGCTTGACGTATTATTGTCTTTGATTCAAATAATATGTTTCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVQLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIIITAAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHISKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSSGGLVCQERRLWKLVGATSFQIGCAEVNKPQVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCACGACACACGCGGGGAACTTTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGCGCGCGCGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGTCGCGA
CGCCCGGCCCCGGCTCGGCGCCCGCTGGGATGGTGCAGCGCTCGCCGCGGGCCCGAGAGCTGCTGCACTGAAG
GCCGCGACGATGCGCAGCGCGCCGCTGCCCCGTGTCCCCGCGCCGCGCCCTCCTGCTCGCCCTGGCCGGTGTCT
GCTCGCGCCCTGCGAGGCCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCACTGCCCTCTGT
TCGGAGTGGGGACCTCTGGATCCCACTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTGAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAAATGA
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGATCATGTGGATCACATCACAACACACCAAACTCGCTGCAAGAATGTGTTTCCACCACCCTCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGTATCGTGGCAGACAACCG
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAGAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTGACAA
GTTTTACAGACCACTGAACATTCGGATCGTGTGGTAGGCGTGGAAGTGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCAATTCACAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTTGTGAGTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
CACGGCAGACCAGTCTGGGGGAATTTGTCATGGACCATTAGACAATCCCTTGGTGCAGCCGTGACCCCTGGCACA
TGAGCTGGGCCACAATTTGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTGAGAA
AGGAGGCTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTCCTGTTTAACTGCGGGAAGTCAGGGAGTCTTTTCGGGGGCCA
GAAGTGTGGGAACAGATTTGTGGAAGAGGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCACCTGTACCTTGAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC
TCACTGCCCAGCCAAATGTGTACCTGCACGATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGTACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA
GAGAGTCAATTTCTCAGGTGATCCTTATGGCAACTGTGGCAAGTCTCGAAGAGTTCCCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCAATTGGTACCAATGCCGTTTC
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGATTCTGTGCCGGGGGACCCACGTGTACTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAACAGGAAGAA
CTGCCACTGCGAGGCCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCC
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTG
CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCACGTCTGCAGAAA
CAGTGCAAGGAAGGGCAGCGACTTCTGGTTGAGCTTCTGCTAAAAATGGAATGCTTTCAGTGTGCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC
ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAAGTGTCCCTTTCCCAAGTGACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTATTATATGAAAAT
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAAACTGAATGGCAA
TGAAACAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC
AGTACTCAGGTTTGGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCTTTCATTTAACAAGTAAGAA
TGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGLWIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCIYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDFK
YRPLNIRIVLVGVEVWMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICFERVNSAGDPYGN
CGKVS KSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLGATKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGP IRQAEARQEAAESNRERGGQEPVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA
ACAAAACAGTTTTTGGGGTTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCAGTACGCTCCACAGCAGTGGGGGAGAAGCAAGGGTT
GGGCCAGTGTCCTCTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTTGGTAGAGATGGGA
TTTCAACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTTTAAGAAGTTAAT
GAAACCATACTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTTCGAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCTT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCATACTAAAACCGAATCCAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCTGTCTCTACTAAAAATACAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVLLAL
LHLYH

Signal peptide:

amino acids 15-27

"04201"=42026001

FIGURE 34

GCCGCGGCGAGAGCGCGCCCGCCCGCCGCGATGCCCCGCGCGCCAGGACGCCTCCTCCCGCTGCTGGCCCCGGC
CGGCGGCCCCGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGCGCTGGGGCGCCCCGGGCCCCAGG
AGGCGGCGGCGGCGGCGGCGGAGCGGGCCCCCGCGGCAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCAACC
TGACACGCGCCGACATGTTTACGCGACGGGATCCAGAGCGCGCGCACTTCGTATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGCTGCAGCCGACTTGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCAGGGGGTGCAGGATACCCACCTTAA
AGCTTTTCAAGCCAGGCCAAGAGCTGTGAAGTACCAGGGTCTTCGGGACTTCCAGACACTGGAAAACTGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGAACCGCCCACTGCCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCTTGAACATTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGGAACACAGGTTCTGGCTATCCCACTC
TTCTCTGGTTCCGAGATGGGAAAAAGTGGATCAGTACAAGGGAAGCGGGATTGGAGTCACTGAGGGAGTACG
TGGAGTCGCGAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCGGTGCTGGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAAATTTCGATGACACCATTCAGAGAAGGAA
TAACCTTCATCAAGTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCCTACTTGGGAGGAACTCTCTA
AAAAGGAATTCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA
AGTATTCGGTACGAGGCTACCCACGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACTTTAGGAACACAGTTGGAGGTCAC
CTCTCCTGCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCCACTGGTGGCT
GTTTCAAGCAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTTCAAGCCAACACACTCTACAG
ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACCTCATGGTCACTGTGTAAACATTTTTCAGTGGCGATA
TATCCCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTCTTTTGAGACTAAAATAGCGTTGAGGGAAATGAAA
TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC
CCACGAGTTCTGGAAAGGTGGCCCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCTGAGTAAGAGAAATGCTGATGCCATAACTTTATGTGTGATACTTGTCAAATCAGTTACTGTTTCAGGGGAT
CCTTCTGTTTTCTCACGGGGTGAAACATGTCTTTAGTTTCTCATGTTAACACGAAGCCAGAGCCACATGAACTGT
TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATTCCACGAGGCTCATTCTCAGTATCTCATTAACCTATTGA
AAGATTCCAGTTGATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCCTTTTCAGAGG
CTTGCTATACTTGGTCTGCTTCAAGGAGGTGCACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC
CAAAGACAGATGTCAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA
ACTCATGTCTGCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCCATATAAACCCTTTGCATCCAACACTCTTACCCACCT
CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCATTATA
CTGTCTGTCTGAGGCAGAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCAG
TATGGTTCACAGATAATCTTTTTTTAAAAAAACCCAACTCCTAGAGAAGCACAACCTGTCAAGAGTCTTGTACA
CACAACCTTCAGCTTTGCATCACGAGTCTTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLGHGGGGRWGARAQEAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLNWMQLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFEHVAQGDHFIFKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESREYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

bioRxiv preprint doi: <https://doi.org/10.1101/130201>; this version posted April 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 36

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATT CAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTT CACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTTCAATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLQLQISLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGCRY
RKFBVGQLDISIARLKTSMKFVKNVIAIGIRETEEFYIIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQYNSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

protein = 43027001

FIGURE 38

GGTTCATATCGATTCTGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCCGTGTGGAGTGGTGGGGCCCTGGGTGGGAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTTT
CGGAAGGGAGGATCAGGGATGTTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCCAGCAGCGCGCCGCGC
ACACCTTTTCTCATTACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCTTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTACCTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTTCTGTGGCTCTG
GTTGCGGCTGGCCAAGGCCGCGCTGCGCACTGCCTTTGTGCCACCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCGAGCTGCGGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCCAGGGCCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCCAC
TCTACCACATGTCCGCTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCACTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGGCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTGCTATGATGTCAACACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGG
CCGTTGCAGGTGTACTGGGCTGTGAGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTAPWKEKSQLERAAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLLPL
LLLKLHLWPQLRWLPADLAFVRLCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRFSYSEAERESNRAARAFLRALGWDWGPDGGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGPIRDPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQ GK
LLKDVFVRPGDVFFNTGDLLVCDDQGFRLRFHRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLS DPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGAGCCGGCGCGCTCCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACCTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTTAAACACTTCTTTTCCTTCTCTTCTCGTTTTTGATTGCACCGTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACCATGCCTTCTTGATCGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTCAACCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCCTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTGAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTCAAGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTCAGCCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTGCGGGGCTTCCCCTGTGAGGCCCTGCA
ACAACTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTTCAGGGATGTGGTCA3CCCAAACCTGCTCCAGCC
CTCAGATCTGCCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCAGGTTGGATGTGGACA
TCACTCGGCCTGACACTTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCAGCCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTTGGTCAAGTGAACCTGATTTTATGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAACCTTACCGTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTCTCATTTAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKA RSCGEVRQAYGAKGFS LADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF
VRTYGMLYMQNSEVFQDLFTTELKRYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFGCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTFEFEVTTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

bioRxiv preprint doi: <https://doi.org/10.1101/140404>; this version posted April 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAAACGCACG
CAGTTTGCAGCGCTGCGCCGGGTGCGCCAACCTACGCAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCCTCACTTCGCCGCTGGTGAGTGTGCGGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTGAGGGCGCTCTGGCCACGA
AAAGTTCTGTCCACTGTGATTCCTCAATTCTTGCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTCTTTTTTTTTCTTGGTGGAAGCTGCTCTAGGGAGGGGGGAGGAGGAGAGAAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTACCCTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA
GGGGCTGTGCGGGGCATCCGCTCCGCTTCTCCACAGGCCTGTGTCTGCTCTGGAAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTGGGGCCAGGCCTTAGAAGAGGAGG
AAGAAGGGGCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA
TTTTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGGTACCACGGATCTGAGATTAAAAACACTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAGTATCAGATACACACCGGACTTCAACATTCATCATAAGACCTACCCAACCAACTGTTTAC
CTCTGGACAATGCCACCCTACCTCAGAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGAAAAATGGCACT
TGGGTTTTTAACAGAAAAGAATGCATGCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCTTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCATA
ACCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTTCATTACCACTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACCTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGATAACATTGACCCCTATACACCAAGGC
AAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGGTCAGGTATCCCCCAAGACCCAGAAGTAACCCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGCAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAGCAAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCACACTGAAAACGTCTCTGCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTGAGTTCCTGCTGTGCTTCACTGCAACTGACCAAACTGCTTTGAATTATAGGAGGAGAACATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAACTACCTTTGATAAATTAC

[illegible][illegible]

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYSWQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLTDKLAAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGVKGWHLGFNRKEC
MPTRRGFDTFFGSLLGSGDYIYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTYGFFYNNSIIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
T TAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTCAGCCT
GGGTCTGTCACTATGGAACATAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAAACAAAT
GCAGATGCTTTCCAGGATACACCGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACTCTAGGACATGTGCCATGATAA
ACTGTCACTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGTCTATCCCTGAAAATTCCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAATAATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACCTTGACGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAC
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAAAGAAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGAAGTGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCAACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAATTCAGTTGTATCAAGGAACTGATGC
TACCAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTTAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTGAGTTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAACTTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGATATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE
ATCEPGCKFGEVCVGNKCRCPGYTGKTCSDVNECGMKPRPCQHRCVNTHGSKYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPK
VNLQPFNYEEIVSRGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT
TGTTTTGCCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGAATGTCCCAGCGTTCTCTGTGC
ATGGATAACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAATC
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTAAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTGCGCGCTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCTACTGCTGCCCCTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCTTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCCGCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTTCAAGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGAGGAGGCGGTGCTAACTG
CTGCCCCTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAAGTGCTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG
CATGAGGTGAGGGGCACATGGTTTCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTCACCGCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTTCGCCGAGGAACCAAGAGCCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCCTGTCTCCCCACCCTGTATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTTCAATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTTTGGGGGGCAGCAGTTTTCTTTTTTTAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRLSNPARPGMLCGGPQPGVQGPCQGDSSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVPVTLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEPE
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGCCCCCATTCGGGCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGGCGGGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCACGACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCGTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCCGGCAGCGCTGCCGTAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCAGGGTGTGCCCCCTGACTGGCAACGTGACCACGTGCGCAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCACAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTACCATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC
TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA
CTGACTGACCCCTGCCTTATTTACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC
TCCAATGGCCGTGATACTAGTGATCATGTTTCCAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC
CCCTCCTTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCTGGC
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTTGTAACCTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVARNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGELEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 51

FIGURE 52

TTCTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACCTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTTCGCTGCTTTTGTAGGGACTTCTTTCCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTCTGTCACCTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAATTTCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTAAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTATACAATGTTCTTTCTTGCTTTGTTTTTATTTTATATATTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTCACTTTTACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194

><subunit 1 of 1, 211 aa, 1 stop

><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL
FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKKGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCTCTCCCGTAGCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC
TGAGGAGATGTTCTCCTCCAGTTCGCGCATGAAGATCATTAACTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTGGGGGGCTTCTGGCTGTGGTTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCAGGGCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCCACCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGGTAGA
GGAGGGGGCCGCTGTACCTGCCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTCTTTGGGGCTTCAGTGAGAACTCTCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAAGTGAATAGTAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

<subunit 1 of 1, 215 aa, 1 stop

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNLVNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLRFQDRVEFSGNPSKYDVSVMRLNVQPEDGEI
YNCYIMNPPDRHRGHGKIHLLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Signal peptide:

Transmembrane domain:

Immunoglobulin-like fold:

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTAGCATCCAACCATCCTCCCTTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC
TCAATGGCTCTGACGCCCCGCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACCTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCAACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCCTCAACGTCCTCAATGGCTTTGACGCCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCTGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCC

FIGURE 58

TGCGGCGACCGTCTGTACACCAATGGGCCTCCACCTCCGCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGACGTCAACCCCAAGTGGTGTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACCTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAAGCTGCTGCTGCTGCTCATCATTGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCTGGACCCAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTGC
TGTTTGGCCACAGTATGGGCAACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGAGGCCTGGAAGGACAAGT
ATATCCGGGCCTTCTGTCTCACTGGGTGCGCCCTGGGGGGGCGTGGCCAAGACCCTGCGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCCAGTCACTCGGGCCCCCTGAAGATCCGGGAGCAGCAGCGGTGAGCTGTCTCCACAGCTGGC
TGCTGCCCTACAACCTACACATGGTCACTGAGAAGGTGTTCTGTGCAGACACCCACAATCACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCACTGCCTCTATGGTACTGGCGTCCCCACACAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACCTGAAGAGTGCCC
TGCACTGCCAGGCTGGCAGAGCCGCCAGGAGCACCAGTGTGCTGCAGGAGCTGCCAGGCAGGCAGCACATCG
AGATGCTGGCCAAACGCCACCCCTGGCCTATCTGAAACGTGTGCTCCTTGGGCCCCGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTGATGGCCACGCGTTTTGCAAAGTTTGTGA
CTCACCATTCAAGGCCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGTGTTTTGTTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGG
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTGGTCCCAGTCCCTGCCCTGGGGCCATG
TGTCCTCCCTATTCCTGTGGGCTTTTCATACTTGCCCTACTGGGCCCTGGCCCCGAGCCTTCCCTATGAGGGATGTT
ACTGGGCTGTGGTCCCTGTACCCAGAGGTCCCAGGGATCGGCTCCTGGCCCCCTCGGGTGACCCCTCCACACACCA
GCCACAGATAGGCCTGCCACTGGTCACTGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGAGTTTGTGCGTTCTTCTGTGGTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGTG
CCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTACCCCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCACCATCA
CACTGCCACCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCAGCTTAGTGCTGGGACTAGCCAGAACTTGAATGGGACCCTGAGAGAGCCAGGGTCCCC
TGAGGCCCCCTAGGGGCTTTCTGTCTGCCCCAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT
GGGGTTCCCAAAGACGCCCTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTTCTGTGCAGCTGGATTTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCCTGGCAATAAAAGTACTCTGGATGCTGTAAAAA

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLLLMLLADPALPAGRHPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSVSTSWLLPYNWTWSPEKVVFVQTPTINYTLRDYRKFFQDIGFEDGWL
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTGAGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAAGTGCCA
GCCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGCGGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTTAGCTAGTGTTTTTTCCTCGCTTTTAATGA
CCTCAGCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGGCCGGGTGAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCCTCA
CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG
CTCCCACCCCTGGCAGCAGGGAAGGGCTTTGCCCTGACAACACCAGCTTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTC
TGTGCCGAGTGATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEYGSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLIVIGDLLFSALWTFLWFGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 61

FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
CCCGTGGCCGAGGCCCGCCAGGTGGCTGGCGGGCAGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCCGGGGTACCTCCGCCTGGTGCCCTGTTTGTGCTGCTG
GCCCTGCTCGTGCTGGCTTCCGCGGGGGTGCTACTCTGGTATTTCTTAGGGTACAAGGCGGAGGTGATGGTCAGC
CAGGTGTAAGCAGGCTGCTGCGTGTACTCAATCGCCACTTCTCCCAGGATCTTACCCGCGGGGAATCTAGTGCC
TTCCGCACTGAAACCGCCAAAGCCAGAAGATGCTCAAGGAGTCTATCACCAGCACCCGCTGGGAACCTTACTAC
AACTCCAGCTCCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTTCATTCTCCAAATCCCCGAG
CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGAGGCACTGCTGGTGAGGAGCTGCTGTCCACAGTCAACAGC
TCGGCTGCCGTCCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
GACATAGCTGCATTGAATTCCACGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG
AAGGGCCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACTCCGG
CTGGAGTGAGCGCTGGCAGAGTGCCGGGACCGACTGGCCATGTATGACGTGGCCGGGGCCCTGGAGAAGAGGCTC
ATCAGCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTCCGGGGGCCATCATGGCGGTC
GTCTGGAAGAAGGGCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCTGT
GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCAGGGCGTCTCAGCACCCCGTACTTCCCCAGCTACTAC
TCGCCCCAAACCCACTGCTCCTGGCACCTCACGCTGCCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGTATGCC
TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGT
GGCTTGCGCATCCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC
TCCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACAGTCCGACCCCTGCCCTGGA
GAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCCTGTGATGGGGTCAAGGACTGCCCCAACGGCCTGGAT
GAGAGAACTGCGTTTTCAGAGGCCACATTCCAGTGCAGAGGACAGCATGCTACTGCCCCAAGGTCTGT
GATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCC
TTCCAGTGTGAGGACCGGAGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGGCCGACTGCAGGGACGGC
TCGGATGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGCCGATTGTTGGTGGAGCTGTGTCTCCGAG
GGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTCCGACACATCTGTGGGGGGGCCCTCATCGCTGACCGC
TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTTCCTGGGC
AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCAC
GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACACCCGGTGGTGCCTCGGCCGCGCTGCGC
CCCGTCTGCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCTGCAGTGTGGATTACGGGCTGGGGCGCCTTG
CGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGCAGCGAG
GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCTGTGAGGGT
GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCCTGGCGGGGCTGGTCACTGGGGCCTG
GGCTGTGGCCGGCCTAACTACTTCCGCGCTCTACACCCGATCAGAGGTGATCAGCTGGATCCAGCAAGTGGTG
ACCTGAGGAACCTGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGCAACTGCCAAGCAGG
GGGACAAGTATCTGGCGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC
CTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGACGTCCCTGAGGACC
CAGGCCACACCCAGCCCTTCTGCCTCCCAATTCTCTCTCCTCCGTCCCCTTCTCCACTGCTGCCAATGCAAG
GCAGTGGCTCAGCAGCAAGAATGCTGGTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
CTGTTTGGGCAGCCTTGCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCTGGTCTAACTTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGGAGGGGACCCCTCAGAGCCCTGGAGACTGCCAGGTGGGCCTGCTGCCACTGTAAAGCCAA
AAGGTGGGGAAGTCTGACTCCAGGGTCTTGCCCCACCCCTGCCTGCCACCTGGGCCCTCACAGCCAGACCCCT
CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTR LGT
YNNSSSVYSFGEGPLTCFFWFILQIPEHRRMLLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGVRLRLKGPDLHLASSCLWHLQGP KDLML
KLRLEWTLAECRDLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGV LSTPYFSPSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDL PCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVC DGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPD CRDGSDEEHCD
CGLQG PSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADR WVITA AHCFQEDSMASTVL
WTVFLGKVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSSGGLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCCGGTGTGCGGGTGCACTATGGCTTGTAACAACAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAACTGCGTTTGAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCTTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAGGGTGA CT CAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG
CTCCGTGCCGCCAAGTTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTTGTTTTTAAA
ACTTCTGTTTTCTTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTGCGCCACAG
CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTTGAAAACCTGGCTGCACAACTCGTGTCAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG
AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC
CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG
TGGATGTCCTCCACACCTACACGCGTTCTTTCGGCTTGAGCATTTGGTATTAGATGCCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTTGCCTTCCAGTGCCT
GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCCTTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL
PLAHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCGGAGCCACCTCTTCCCCCTCCCCCGC
TTCCCTGTGCGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGGCTGACAGT
CGGCAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGGCGCTCG
CTGCCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGGCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGGCGCGGCGGACGGAGAAAACAACCTCCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG
CCGCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTGAGACCCGGGCCATGGAGCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTCCG
GCACCTCTGGACAGCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTAGAAGTGCAGGGCACCTTACAGA
GGCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCTGGGCAGCAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCTACGCTCCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCAGCCCTCTGCAGCTGCCCGGGGCAACGTACCATCACTTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCTGCTCTCTACAGCCAAGATTGGCTGATGTGCTGCAGGAAG
AGTTTCAGTGCTGAACACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCCCCTCCCTGCCTTGCAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCTGGATATACACACCTAGCCTCAGTCTCCACCCCCAGT
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGTATTTGCTTGCCTTGGGACAGACCCTGTGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACCGCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGTGGCAGCT
CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGCGCATTTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTTCGACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
CTTCTCTACGGGCAGCTCATTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTACAGAGAATCCTAATGATA
ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCCAGGTG
CCCGCCGTGCTCAGCGGGGCGGCTTGATGCGACGCGCTGGTACGCCGTCTCCGCCGTGGGGCTTGCTCCCTCGAA
CCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTCACACCTTCTGCTGCTCCCCCTGAGGCGCTAGATG
GTGGCACAGGTCCAGCCCGTGAGGGCGGGCAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCA
AGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGACGGCCCTGCGAGGCCGCTGTTGCCCAGCCTGGGGCCCC
CAGGACCAACCCGAGCCCCCTGGACCCACACAGCAGTCTTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCCTCTCCCCCTGGGGGCTCTACTCATAGTGGCACAACCTTTTAGAGGTGGGTGAGCCTCCCCCTCC
ACCCTTCCCTTCCCTGTCCCTGGATTTAGGGACTTGGTGGGCCTCCCGTTGACCCCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG
CTCACAAAAAGAGTGCACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAAATAAA
GGAATCATACATCTC

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTQLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEETFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPOQSWLLDPHDGRRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDGADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPCADGSDEWDGS
YVLPKRVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQGRRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLPSLGPFGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVVVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC
TGAACCATATATTGTTATCACTGGATTGGAAGTCACCGTTATCTTATTTTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTNTTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTIONTTGACACAGAGAGGT
CTAGCACATAACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTC
CAATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTTCGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA
GAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCCTGACCTCAAGCAA
CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGCTTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVYKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCCTGTGGCCGGACCCGTGAGGCGTTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACCTGTGCGGGACGTCTTGACTTCAAACGCCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCCCTGCGCCATGTTCCCTGGATGGCTGCGCCCACCTTTTG
CGCCCATTTGGCTTGGCTGGTGTCTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACCTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCGGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLTHLLLPCLKACAPSRVVVVASAAHCRGRLDKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHGPVNSELFLRHVPGWLRPLLRL
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDDDRAAHLWEASKRL
AGLGPGEAEPEDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPOSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATCATGGGAGGCATGGCTCAG
GACTCCCCCGCCCCAGATCCTAGTCCACCCCCAGGACAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCACCATCCGCTGGTGTGCTGAATGGGCAGCCCTGAGCATGGTGCCCCCAGAC
CCACACCACCTCCTGCCTGATGGGACCCCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCACGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGTGGCTGTCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCACAGTCTCATGGTGGAAAGATGGGAAACCCCTG
GCCCTCCAGCCCCGAAGGCACACAGTGTCCGGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
ACCTACATGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCCGGGTTTCCATCCAGGAGCCC
CAGGACTACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCAGCTGGAAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTGAGTGGCCCTGCTGCGCCTGCCCAA
TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCC
GGCTGGCAGAGCGCAGAGCTTGGAGGCCTCCACTGGGGCCAAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCCTGAGGCTGCCGGAAGAGTGGCCAGTGCCCCACCTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCCACCACCTGCTGAAAACCAATGGCATC
ATCCGTGGCTACCAAGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAAGTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGGAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAAGTGGCTGCAGTCACTGGTGTGGAGCT
GGGGAGCCAGTAGACCTGTCTGCCTCCTTTTAGAGCAGGCCATGGAGCGAGCCACCCAAGAACCAGTGAGCAT
GGTCCCTGGACCCTGGAGCAGCTGAGGGCTACCTTGAAGCGCCTGAGGTCAATTGCCACCTGCGGTGTTGCACTC
TGGCTGCTGCTTCTGGGCACCGCCGTGTGTATCCACCGCCGCGCGGAGCTAGGGTGACCTGGGGCCAGGTCTG
TACAGATATACCAAGTGTGAGGATGCCATCCTAAACACAGGATGGATCACAGTGAATCCAGTGGTTGGCAGACACT
TGGCGTTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCAGCCTCAGCAGTCCGGCTGGGGGCGGATGCCCG
GACCCACTAGACTGTCTGCTGCTCCTTGGCTCCTTGGGACTCCCGAAGCCCCGGCGTGGCCCTGCTTCCAGACAC
AGCACTTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCAGCCAGGCCAAGTCCCCAGGTCCCAGCT
GTCAGGCGCCTCCCACCCAGCTGGCCAGCTCTCCAGCCCTGTTCAGCTCAGACAGCCTCTGCAGCCGAGG
GGACTCTCTTCTCCCCGCTTGTCTCTGGCCCCCTGCAGAGGCTTGAAGGCCAAAAGAAGCAGGAGCTGCAGCAT
GCCAACAGTTCCCACTGCTCCGGGGCAGCCACTCCTTGGAGTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCTTTCCCAAAGCCAGGAGCTGTGCCCAAGCTCTGGTTGCCTGGCGGGCCCTGGGACCGGAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCATCTCCCTCCAGCACCCCTCTTTCTCATGAAACTCCCCAACCTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACCACAGGCTCCCTCCTCCATCCTGCTGCCAGCAGCCCCCATCCCCATC
CTTAGCCCCCTGCAGTCCCCCTAGCCCCCAGGCCCTTCCCTCTCTGGCCCCAGCCAGCTTCCAGTGCCTGTCC
AGCTCCTCACTGTCTATCCCTGGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTCAGTGAGGGTGAGGAGACTCCCAAGAACAGCGTCTCTCCATGCCAAGGGCTCCTTACCCCCCACCACCTAT
GGGTACATCAGCGTCCCAACAGCCTCAGAGTTACCGGACATGGGCAGGACTGGAGGAGGGGTGGGGGCCAAGGGG
GGAGTCTTGTGTGCTGCTGCCCCACCTCGGCCCCGCTCACCACCCACCCAGCGAGGGCTCCTTAGCCAATGGTTGGGGC
TCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCTTGTGAGCTCCTCCGATGGCTCCTTCTCTGCTGAT
GCTCACTTTGCCCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCAGGGAGGCAGACTGC
GTCTTTCATAGATGCCTCATCACCTCCCTCCCCACGGGATGAGATCTTCTGACCCCAACCTCTCCCTGCCCCCTG
TGGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTGAGCCACACCCAGCGGCTGGGAAGGGGGATGCCCTCCC
TGGCCCCCTGACTCTCAGATCTCTTCCCAGAGAAGTCAGTCCACTGTCTGATGCCAAGGCTGGTCTTCTCCT
GTAGATTACTCCTGAACCGTGTCCCTGAGACTTCCAGACGGGAATCAGAACCCTTCTCCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
TTCTCCCTCCACGATTGTGAAAACAAATGAAAACAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAAA
ACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCCTCTGTGCCATCCACTCCACCCAGGTTGTTTGGC
CTGAGGAGCAGCCCTGCCTGCTGCTTCCCCCACCATTGGATCACAGGAAGTGGAGGAGCCAGAGGTGCCCTT
GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGAGCCCCCTCTCAGCCTTACCT
GGGCCCCCTCCTTAGAGAAGAGCTCAACTCTCTCCAACTCACCATGGAAGAAAATAATTATGAATGCCACTG
AGGCACTGAGGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAAGCACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAAATTTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGRMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSPAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGWP
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKKQELQHANSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTTGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTTGTGACCCTGTGGA
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTCTGTTTCCAGGACCCTGCTGTCTCTCCCTCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCAGTGGTCCGCATACTGGCCCCAGTCTGGTGCTGCTGAGCCTTCTGTGAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCTCCTCAGGCTCTCCTCTTGATGTTCCA
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCTGGGC
CTCATGCCCAGTGTGCGACCCCTGCCTTCTCCTCCACTCCAGACCCACCTTGTCTTCCCTCCC
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGACG
GACTCTGAATTCTAACAATGCCAGTGACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTTCACTTAGAGTCTGCATTGGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTGAGTGCACATTGCAGGATAAGCCAGGACCGGCACAGAAGTGG
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
GCACAACTACTATTTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTGTTTCAATATGCAAACTTGGAAAG
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTCTGTTTTGTGTAATCTTCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT
TTTACAGAGCAATTATCTTGTATATACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSIRD SRQELSLIVTLWNLTLDAGEYWCGVEKRGPD E SLLISLFV
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQATETQRNEKFWLSRLTAE E KEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTTACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCTTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTNPVNPVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCGGGCGCCCGCGGGCACCATGAGTCCCCGCTCGTGCCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCGGGAACCGGCGCTGGAAGTCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGACAGAGTGGGGAGCT
GGAGAAGTGCGGCTGTGACAGGACAGTGATGGGGTCAGCCACAGGGCTTCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTCAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGCTGCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATTGACCGCCTGCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTTTGTCTGGTTTGGTTTTTGGGTCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGACAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTCTATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAAGAACTCTTAAGTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATACCAGGTCAAGGGCACCAAGGTTCAATTCAGCCCTTACATGGACAGCTAGA
GGTTCGATATCTGTGGGTCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCACCCCTAGAACCCAGCCTGCCCCAGCCTGCCCCCTGGGAAGAGGAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCCGAGTCACCTTTCACAGCGCTGTTCTCCATGAAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCAGTCACCTTTCACAGCACTGTTCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRVQMCKRNLE
VMSVRRGAQLAIEECQYQFRNRWNCSTLDSL PVFGKVVTQGTREAAFVYAISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKT CWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTC NKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCTGGTC
ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTGCGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACCTCGCGAGCCG
TGCCCCACGTCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTACAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCATGATGC
TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACTGTACCCCGCCAGTGCCCTGGAGCCGCGCCCATTCGAGCATGTCGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA
AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGSSEEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAEELGEAQAKLMEQESALR
ELRERVITQGLAEAGRGREDVRTELFRALEAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEPNDAWGRENVCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCCTTCTCCTCCTTTCTCCACGTCCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAAGTGGAGCCTCATTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAACCCAGCCCGGCCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTCAGTAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTTCGGCGCCCCCGCTCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTCTGTTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCAACTTCGCCACCATCCCGCAGGACACGGGTGACCGAGATAACGTCCTCCTCTCCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCTGCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCTGCTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAACTGCGTCTAAGACCAGAGCCCCGAGCCCCTGGGGCCCCCGG
GAGCCATGGGGTGTCGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTTCATCG
TCCAGGGGCCTGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFGVDSLCLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSEFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLP SRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGR LGTKSRTRYVVRVQPA
NNGSPCPELEEEAECVDPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 87

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAATA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACCTTTGTTTGGGAAGCTTATTATGACAATACCATTTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCAATCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTGAGCAGATGAACT
TAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAAATTTAAAGGCTGAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAGCAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAA
TTAAAAAAGGACACAAGTGCGAATGTFAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACTCTTAGCAG
CAAAACAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCCTCCAGATGGTGTGTTGCCGAATACAGAAGAGAAAAGCAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTTCTT
TCCACAAAAA

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919

><subunit 1 of 1, 472 aa, 1 stop

><MW: 53847, pI: 5.75, NX(S/T): 2

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAESEEEFVNRSQSMKGKSKSSHLLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKCLKKDTSANVKSAGEGEVEKKS
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAEPPDGAVAEYRREKQKYEALRK
QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQSDTTFEIIDPRNPVNKRREESKCLMREKKERR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTTCGCTTGCTGACGGCGTCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCTCG
ATGTACGTAGCACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGCTGGTGTCTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACCTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTCCGCCAGTCCGGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTACCCCTAGCCTCGGGTCTCTTCATCGGGC
CATTTCCAGAGTGGCACCGCGTTATTAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCAAC
GAACTTCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCATTTGGCTCTTGCCCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCATGGAAGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCCATGATGCCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT
CCCACCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTCGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCCTGCTCAACCTTGTTCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTTGGTAGTTTGGGA
TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTGATTAGAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATT SAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGT LQGKQMHVGKTPIQVFLGV PFSRPP
LGILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLN VYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVT LFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVS NKMRF LQLNFQRDPEEIIWSMSPVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMD
IVQDATFVYATLQTAHYHRET PMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

[illegible][illegible]

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSF SYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQSPDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins
signature.**

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAACAATTCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCTATCAAAGTTTCGGATATCAATGACAATGAACCAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAACCTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTGCGAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAA
AAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCCCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCACAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAGTTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTTGACTAATAGAACTGGTTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATA
TTTGGGTTTTATTTTTTTGACTTTGGGTTTTAAACAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAAGAGAGAATATATTCCAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTTTACCATCAAAATTTTTTAAAAGTGCTAATGTGTATTGGAACCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTTAAGCTACCTACATGCTGTCTATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAATAATG
TAGGAAGATATTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCATGAAGTATTGTTTCCTTTAT
TTAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLPCLGATENSQTKKVKQPVRSHLRVKGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTGIVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSLSD
VNDNKPIFKESLYRLTVSESAPTGTSGITIMAYDNDIGENAEMDYSIEEDDSQTFDIIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHFHYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

[illegible][illegible]

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCTTAACCTCCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCACCTGCAAACCTCTCCGCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTCATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGAGG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCCAGGCCATGTACGAGGGGCTGTGGA
TGTCTTGCCTGTCGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTCTTGACAGGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAAA
ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCTCA
ATATAGGAGGGAAGATTTTTCCATTTGTATTACTGCTTCCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCTGCC
TTATATACATATGTAACAGTCAAATATCATTACTCTTCTTCATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACCTTTATTGATTGAATTTTAAAGCTACTTATTTCATAGTTTTATATCCCCCTAACT
ACCTTTTTGTTCCTTCCCTTAATTGTATTGTTTTTCCAAGTGTAATTATCATGCGTTTTTA
TATCTTCCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACCTATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCAGTGCCTTCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGTGTTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACCTGCATACGTTTTTGGTGTGCTTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI~~L~~AF~~L~~GWIGAI~~V~~STALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSL~~L~~NLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTPV~~N~~ARYEFGQALFTGWAAASLCLLGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTGCGAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCTGTCCC

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTA
TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTCGCAGANCACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAAGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTGCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCCTGTCCC

[illegible]

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCCTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATAATTTCTTCTTGCAAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTTCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCTTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTTCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTTCCTG

TCATAGGT=430400T

FIGURE 106

TTCCTGGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCCCTGTCCCCGAAAAACAACCTCTTACCCACG

Figure 106

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTCNGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTTGC

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCGCCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAACTTAAAGCGGGCGCAGCATTAAAGCTTCCCCCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCAGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAATAAAATTATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCAATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAAATGACACCCTTCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTTCCTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTGTTATTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTATCTCATGTTTTCTTATTGTCAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTCTGGGTGAGTAAATAACAACCTGTATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCGTGTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHCLKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNNGIID
AGASINVSVMQLQPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

protein: 4304007

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTG GGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

4004-0001-0001-0001

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAACGTATTTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA

FIGURE 113

GGTGGCCCATTCCTCGGCCAGGCTGCTTTCCTGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTTANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTG
TTATTGAGAGATGTTTAAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

48027001

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACGTGTTATTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

430100T

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANCCCC
GGCCCAGGCTGCTTTCGCTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 115

Variable	Mean	SD	Min	Max
Age	38.5	12.5	18	65
Gender	0.5	0.5	0	1
Marital status	0.7	0.5	0	1
Education	12.5	2.5	8	16
Income	1500	500	500	3000
Health status	0.8	0.4	0	1
Smoking status	0.3	0.5	0	1
Alcohol consumption	0.2	0.4	0	1
Exercise frequency	0.5	0.5	0	1
Stress level	0.6	0.5	0	1
Sleep quality	0.7	0.4	0	1
Work satisfaction	0.6	0.5	0	1
Life satisfaction	0.7	0.4	0	1
Depression score	10.5	5.5	0	30
Anxiety score	12.5	6.5	0	35
Quality of life score	75.5	15.5	30	100

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTGAGTGAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTTCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTTCAAGCATCTTGACAAA
AGGTTCTTAACCAATTTCCCTTTTCAAGCAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCTCAGCATGTTCTAACTGCTGCCCCTGTGTTTATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTCT
AGAGGATTGCCGAAGGGAGGCCCTTCTTTTCAAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCCGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAAGTGGGTGGATGTCCACGGGGTTTCAAGAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCCATAAATACGCCCAGATTTGCCTCTGGATTACCG
GAACGATGCCAATTGTGCTTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCACATTTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACCTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGTAGATTTTCAATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTGAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACTAATAACTGTTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTATAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCTTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFAEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAHCVHDGKDYVKGSKKLRV
GLLKMRNKS GGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDSNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCAACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACCTCCCTATGGAAGTAGACATATTCGCCCTTGCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACCTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCTCAGAGGAGCCCTCGTTAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAGPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVKGPDLHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPSC
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPASPAGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQEELEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAAC TCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCC
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGACTCCTCAAAAGTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAAC TGGCTAACATTCTTTT
TACCAGGGAAGTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGT CATGGGCTTTTTC AAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGT CAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAAC TCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAA TAGGAACAAGGAGTAAAAGAGCTGTTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTT CAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAAATTAAGTGA CTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEAAAGQLRREL RQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVS SKLYKYGDINFDDLNSEQSYNKSFCYSR SKLANILFTRELARRLEG TNVTVNVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

protein = 336 aa

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTTGGGATAAGAGAATTTTCAGCAAAGATGTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTTCAAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

[illegible]

ATTAAGGAAGAAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAAACCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGGCTCCCGGAGCCCAGCCCTTTCCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCTTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCACGCGTCCGATGGCGTTTCACGTTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTA

CTCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTTGCTTGTGGAAAGACTG
TTTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAACCTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACTTGTA
GGGCTCATTTTGGTTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTTAGT
GCAATACAATAAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

protein: h3d4t001

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTA CTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 130

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAAGTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSILWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLFPPLTLVRSFWSDMMDSAQSFITSSWTFYQLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 132

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGTTGCCAGAATCAGCTGCCATTCGC
TGAAGTGAAGACAAGAACAATTATGTCCCTGATGCCAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTGCGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

Figure 134

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCCGGAGGT
GGGGCGCCGCTGGGGCGCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCGCGGCTG
GGGATTCTTGT¹TTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAAATATGGAATGTCTACGAAGAAAACCTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCCCTTG
CATTTTGATGAGAATTCATTTTTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAQAQRFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDFRYKVNLRPCPFWNDISQCGRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAF
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

protein = 460/2100

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

[illegible]

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCTTTCGTCACTCACCTGTTCTTGCCCCCTGGTGTTCTTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCTCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTCT
ATGGTGAGCTTAAGGAGAGGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAAGAGAA
GTGTGGTAAGGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACATTTCCCCCATCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCCTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

protein = "protein"

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCCTTCCAACCTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAAGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACCTCACCTGTTCTTG
CCCCTGGTGTTCCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACCTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 140

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTTCAACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTGCTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCTGGTACTGGCCCTGTTTGCCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTTCGTGCCACTGTTTCGTCTGGAATAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCCTCCAGACACCTTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTGCAGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGAAGTGGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTACAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAA
AAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLSFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLIMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAEHV KMRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFM LILV
VVPLFVWKMGRL LQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

biochem 4304001

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTTGGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTAACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCCTCCTTTACAGCCACCGACTCAGTGAAGTGC GGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTTCACTTCAACCAGGAACTCTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGAGACTCCCCCTTCGAGGATTGCACCCGCCCGTCTTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLFLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSLRLLSQNPFSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTGCGCCGCCAGCCTCCGCCGCCGAGCCTC
GTTTCGTGTCCCCGCCCTCGTCTCTGCAGTACTGCTCAGAAACGCTGGGGCGCCCCACCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCAACCCCACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGCGTATTCTGTCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCGCGCGCTCCGTTTGGCT
CACCTCTCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
ATCCTGAGGTCATTTCATTATGAAGTGTACCGCGCGGAGTGGCTCAGAGTAACCACAGTGTCTTTCATGGCTAGA
GCAATTCCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAAACAACGAGGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAGA
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGGAAGATATAGGCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTCGTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCTATAACATGAACATCTGGGGGCG
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAACCTGGTGGGGCCATGCCCTTACAAA
CATGGGCGGCCCTGTTCTGCTTGCCACCTAGTTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAAACAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCCGACAAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCTGCTGGCTGTTTGGAT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTG
ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCT
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTGCAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
AAGACCTACATTGCTTCTTTTCAAGATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
AGAGTGTTTGCTGTTGTGGAAGTGAATACTTGGAAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
ATTTTGTATAAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCAAAAGGTGCCAAATGCATA
TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTTCATATATTATATGGTGTCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCC
AGTTCAATGCTATGCCATTACCAACTCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
TGTTAATTTAGGCATATAGAATATTAAATTCTGATATTGCACCTTCTTATTTTATATAAAATAATCCTTTAATATC
CAAATGAATCTGTTAAATGTTTGATTCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
ATGAAAACATTCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
AAATTGAGGTACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
GAACAAAGATGAACATAATGTATTACATTGCCACTGATTTTCTTAAATGGTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGATATATTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
AATCTGTAAATGTTAGTTTTTGGTAATTTTTTTCTGCTGGTGGATTACATATTAAATTTTTTCTGCTGGTGG
TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAIAMVVPNATLLEKLLEKYMDEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVWATSNRIGCAINLC
HNMNIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

TECHNICAL REPORT

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCCGCGCTCCGACGGGCCAGCGCCCTCCCCATGTCCCTGCTCCACGCCG
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTGAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGCTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCT
GGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTAAGTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 148

FIGURE 149

MSLLPRRAPPVSMRLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMWIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

protein database

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMPLIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTSLHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLG
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEV
FGIPMIGPTTELFSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPQDTKFIHTKPNRFEE
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPDM
FPYGTRRSPAKIWPTTKRPAITPANNPKHSDPHKTGPEDTTVLIETKRDYSTELSVTI
AVGASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLE
HDHECESLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNT
FSGGQNSTNLPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCAGTACTTGAAACGGGAGCACTCGCTGTGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCTGTTGATTGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAACCTGACAGTGGAGAGAACCCCAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTCTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCTCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCAGGCTGCCGTGTTGTTTACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTGCTTGGGTT
TTGCATTTGACCCAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAACT
TCTTCCCTGCCTTACCTTCTTTTCACTCCATTCAATTGTCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGTTGGTGTAATAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCAATTGAGCAGCACAGCCTGAGTGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPPQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGalWNRVPCFLRDWELQVHFkiHGQgkkn
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDtYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTelGGCTaIVRNlHYDTFLVIRYVkrHlTIMMDIDGkHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLklFELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

bioRxiv preprint doi: <https://doi.org/10.1101/123456>; this version posted January 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCTGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCAATCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAAATCTGTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCCGACCCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTGATGCCC
CGCATCATTTTAAAGCTGGTCTCAGTGCATGTGTTTCATTCGCCACGGAGACAGGTACCCACTGTATGTATTCCCA
AAACAAAGCGACCAGAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTCA
TTAGTCACATGTCAAAGGATCCGGAGCCTCTTTGAAAGCCCCCTTGAACCTCTTGCCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTCAGAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCCTGGGAAAA
GCCGGACCCCTACAAAGTGGGCTGGCCCTTGCTTTATGGCTTTCTCCAGATTTTGAAGTGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCCTGTTCTGCTCTGGAAGCTGCTATTGCCCCGTAAGAAACAGTATCTGGAAAAGGAGC
AGCGTCGTGAGTACCTCCTACGTTTGAAGAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTGAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGACTTCGGGTATTCTCTCTGGGTGCCACCCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCCTTGGGCCCTTTCAGAAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCAGTGAACATTCCGTCCGGATTCTTTACAATGGCGTCGATGTCAATTCCACACCT
CTTTCTGCCAAGACCACCAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAGAACTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTACAGGGAAGGATTTCTAAAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATTTGCTTTTAAAGGCTAAATATTGTTTGTGGGAACACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTTCAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACAGCACACCTCAACCAAATTTTTTTAATCTTAGACATT
TTTACCTTGTCTTGTGTTAAGAATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTCTGTAAAGG
GCCAGATTGTAATATTTTCACTGTGTGGACCAAAGGCCACATACAGTCTCTGTCTATACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCCCTGGGCTGTAGTTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAAATGTC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCCCTCCAGTAACCTTCTGCTAGAAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAAATATT
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAACTAATTATGGCTCATTCCTTTGACAAGCTGTAGA
ACTGGATTCTATTTTTAAACCATTTTTCATCAGTTTCAAATGGTAAATTTCTGATTGATTTTTAAATGCGTTTTTGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACTAATGGTGCTAATTACAGAGAAATGGAAGTGAAAGTGAGATTCTCTGTTGTCATCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTTGCATTTGAATATGTCTGTTTCTATAAAATAAATTTTTTAAGAATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPPSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLNLVRFVKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 155

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA
 GCTATTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
 GCCGCTGTTTACCAATCGGGGAGAGAAAAAGCGGAGATCCTGCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGAT
 AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT
 GGGCGAAGGAAGTCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT
 TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
 ATGCACCCAGATACGCTATTCACTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
 GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCGAGCTTTTCGCCCT
 GAATCCGCGCAGCGGCAGCTTGGTCAACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
 TCAATTAAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
 CGACAATGCGCCTTACTTTCTGTAAGTGAATTAGAAAATAAAATAGTGAAGTGCAGCCACTGAGATGCGGTT
 CCCTCTACCCACGCTGGGATCCGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGAACACTCA
 CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCCGAATTGGTGCTGAAACGCGCCCTGGACCG
 CGAAGAAAAGGCTGCTCACCACCTGGTCTTACGGCCTCCGACGGGGCGACCCGGTGCACAGGCACCGCGCG
 CATCCGCGTGATGGTTCTGGATGCGAACGACAACGACAGCGTTTGTCTCAGCCCGAGTACCGCGCGAGCGTTCC
 GGAGAATCTGGCCTTGGGCACGCGAGCTGCTTGTAGTCAACGCTACCGACCCTGACGAAGGAGTCAATGCGGAAGT
 GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCAAGTTTTCAAAGTAGATTGTAATTCAGGGACAATATC
 AACAATAGGGGAGTTGGACCACGAGGAGTCAAGATTCTACCAGATGGAAGTGAAGCAATGGATAATGCAGGATA
 TTCTGCGCGAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCTCACCTCTCT
 CGCCAGCTCGGTTCCCGAAAATCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA
 GGAAAACGGACAGGTGATCTGTTTCAATCAAGGAAATCTGCCCTTTAAATTAGAAAATCTTACGGAAATTAATA
 TAGTTTAGTTCACAGACATAGTCTTGGATAGGGAAACAGGTTTCTAGCTACAACATCAGTGACCGCCACTGACCG
 GGGAAACCCCGCCCTATCCAGGAAACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCCGGCTCTT
 CCCTCAGGCCTCCTATTCCGCTTATATCCAGAGAACAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
 CGACCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
 GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
 CTTGCAAGTGAAAAGTATGGCGCGGGACAACGGGCACCCGCCCTCAGCAGCAACGTGTCGTTGAGCCTGTTCTG
 GCTGGACCAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCCTCCCAACGACGCTTCCACTGGCGTGAGCT
 GGCTCCCGCTCCGACAGCCCGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
 CTGGCTGTCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
 CACGGCGGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCACGGCCAGCC
 CCTCTCTCCGCCACTGTCAAGCTCACCCTGGCCGTGGCCGACAGCATCCCCAAGTCTTGGCGGACCTCGGCAG
 CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG
 CGTCTTCTGGCCTTCTGTCATCTTGTCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTACGCTGCTGAGGC
 TTCAGGAGGCGGCTTGCAGGAGCGCGGCGTGCACCTTTGTGGGCGTGGACGGGGTGCAGGCTTTCTGTCAGAC
 CTATTTCCACGAGGTTTTCCCTCACCACGAGCTCGCGGAAGAGTCACTGATCTTCCCCAGCCCAACTATGCAGA
 CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAGCGAGCCCTTTTGTCTGTCAGGTGATTCGGTATTTTCTAAAGA
 CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTAAATTGCTCTGTCTCCCAAGC
 TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAATCTTAGGCTCAAGCAATTATCCACCTTTGCCT
 CCGGTGTAAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT
 CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCACGCCTGTAATCCAGTACTTTGGGAGGC
 CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAACACAGCCCTGACCAACATGGAGAAACCCCGTCTATACTAA
 AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT
 TGCTTTAACTGGGAGGTGGAGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG
 AAATCTATCTCA

FIGURE 157

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVP EELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFP LPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRY SFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVS LVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILL LALRLRRWHKSRL LQASGGGLTGAPASHFVGVD
GVQAFLQTY SHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSV SQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

protein: h001001

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCTTT
ACAGCTGCCCCACTGTAAAAAACCAGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 158

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGACACGGGCGGGCTGAGCGCCTCCTGCGGCCCGGGCTGCGCGCCCCGGCCCCG
CGCGCCGCCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCCGCGCCCCG
GCCCCGCGCCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCGCCCCGGCCCCCGCCCCGCCCCG
CCCCGGCCGGCGGGGGAACGGGCGGATTCTCTGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCGGGGCCGGGCGGTAGCGGCGGCGCCTGGA
TGCGGACCCGCGCCGCGGGGAGACGGGCGCCCCGCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCCCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCACGGCCTGGGCCGCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCTACTGCACCAGAACCGCGTGGCCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCTCATGACACTCTATCTGTTTGCCAAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGGAAGTGGTGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCTGCCGCCCGCCACCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAA
AAAA

FIGURE 160

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRLLAWLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHG NRISHVPAASFRACRNLTILWLHSNVLARIDAAAF TGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRRLNDNPWVCD CRARPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA DKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVS RGNKHLQLNQETADLLLNEKLDREDLCGHTEPCVLR FQVLLES PF EFFQ AELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLVDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATD VDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDH APEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISC SIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAF TQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELG LFGVWAH
NGEVRTARLLSERDAAKHLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAA
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFP PQC PGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 163

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCTCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCTTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATAACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCCTTATTCTTTTAAATTAGTTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETCLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLSFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTTA
GTCCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACCTAATA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGLLLLFLASHCCLGSARGLFLFGQPDFS YKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHS LCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLVGVSEKDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGELVITSVKRWQKGQREFKRISR SIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

protein database

FIGURE 168

GTGGAGGCCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGCAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACCCTG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTTGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTATCCTGGCTGTGTGCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCTGA
GTGTTCTCTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTTGGGGCCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCCACCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPLV
KKGEDIPMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQIQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGLIMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPPIPAHCKVLIKVLVDNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLAQAQD
QGLQPLSAKKQLSIQISDINDNAPVFEEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNVEEMAGFEFQVIAEDSGQPMLASSVSVWVSLIDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLNRQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCCA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGA CTCTC
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGA AAATCTTACTCTTATTCAACAGGAAGTGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGT CACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTGGTCCCAACACATT
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGAGAACTGCAGTTCAACTTCTATCACC GTTG GTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAA AATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYAIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLEFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMAALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGMVFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

Variable	Mean	Standard deviation	Minimum	Maximum
Age	34.5	10.5	20	55
Gender	0.5	0.5	0	1
Marital status	0.5	0.5	0	1
Education	12.5	1.5	10	15
Income	1500	500	1000	2500
Health status	0.5	0.5	0	1
Employment status	0.5	0.5	0	1
Home ownership	0.5	0.5	0	1
Vehicle ownership	0.5	0.5	0	1
Life satisfaction	4.5	1.5	1	7
Subjective health	5.5	1.5	1	7
Life expectancy	75	5	60	90
Quality of life	5.5	1.5	1	7
Healthcare utilization	0.5	0.5	0	1
Health insurance	0.5	0.5	0	1
Healthcare expenditure	1000	500	500	2000
Healthcare access	0.5	0.5	0	1
Healthcare quality	0.5	0.5	0	1
Healthcare equity	0.5	0.5	0	1
Healthcare efficiency	0.5	0.5	0	1
Healthcare sustainability	0.5	0.5	0	1
Healthcare innovation	0.5	0.5	0	1
Healthcare leadership	0.5	0.5	0	1
Healthcare governance	0.5	0.5	0	1
Healthcare accountability	0.5	0.5	0	1
Healthcare transparency	0.5	0.5	0	1
Healthcare integrity	0.5	0.5	0	1
Healthcare trust	0.5	0.5	0	1
Healthcare collaboration	0.5	0.5	0	1
Healthcare partnership	0.5	0.5	0	1
Healthcare coalition	0.5	0.5	0	1
Healthcare network	0.5	0.5	0	1
Healthcare system	0.5	0.5	0	1
Healthcare organization	0.5	0.5	0	1
Healthcare management	0.5	0.5	0	1
Healthcare practice	0.5	0.5	0	1
Healthcare research	0.5	0.5	0	1
Healthcare education	0.5	0.5	0	1
Healthcare training	0.5	0.5	0	1
Healthcare development	0.5	0.5	0	1
Healthcare improvement	0.5	0.5	0	1
Healthcare innovation	0.5	0.5	0	1
Healthcare leadership	0.5	0.5	0	1
Healthcare governance	0.5	0.5	0	1
Healthcare accountability	0.5	0.5	0	1
Healthcare transparency	0.5	0.5	0	1
Healthcare integrity	0.5	0.5	0	1
Healthcare trust	0.5	0.5	0	1
Healthcare collaboration	0.5	0.5	0	1
Healthcare partnership	0.5	0.5	0	1
Healthcare coalition	0.5	0.5	0	1
Healthcare network	0.5	0.5	0	1
Healthcare system	0.5	0.5	0	1
Healthcare organization	0.5	0.5	0	1
Healthcare management	0.5	0.5	0	1
Healthcare practice	0.5	0.5	0	1
Healthcare research	0.5	0.5	0	1
Healthcare education	0.5	0.5	0	1
Healthcare training	0.5	0.5	0	1
Healthcare development	0.5	0.5	0	1
Healthcare improvement	0.5	0.5	0	1
Healthcare innovation	0.5	0.5	0	1
Healthcare leadership	0.5	0.5	0	1
Healthcare governance	0.5	0.5	0	1
Healthcare accountability	0.5	0.5	0	1
Healthcare transparency	0.5	0.5	0	1
Healthcare integrity	0.5	0.5	0	1
Healthcare trust	0.5	0.5	0	1
Healthcare collaboration	0.5	0.5	0	1
Healthcare partnership	0.5	0.5	0	1
Healthcare coalition	0.5	0.5	0	1
Healthcare network	0.5	0.5	0	1
Healthcare system	0.5	0.5	0	1
Healthcare organization	0.5	0.5	0	1
Healthcare management	0.5	0.5	0	1
Healthcare practice	0.5	0.5	0	1
Healthcare research	0.5	0.5	0	1
Healthcare education	0.5	0.5	0	1
Healthcare training	0.5	0.5	0	1
Healthcare development	0.5	0.5	0	1
Healthcare improvement	0.5	0.5	0	1
Healthcare innovation	0.5	0.5	0	1
Healthcare leadership	0.5	0.5	0	1
Healthcare governance	0.5	0.5	0	1

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGACCATGTTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTNNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTCCCATTC

48021001

FIGURE 176

CTCGCGCAGGGATCGTCCCATGCGCCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTCGAACCTGCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATCGCACTGGAGGCCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCACTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCTTCTCCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACCG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCCAACCA
CAAGGGTGTGTGGTCACTCTGCGCAAGGACAGCGCCAGTGCCTGGTGGCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCCCTCTCCGGCTCTGCGGCTCCCCCTGACTCCATGTTTCGGGATCAGCCTGGCTGTCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCCCTTTGATGGTGTGGGAAAGTCTTTCATCTACCA
TGGGAGCAGCTGGGGGTGTGCGCAAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGTCTTTCAGGGCCAGACCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCAACTGTGTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCTGAGCCGTAACCTGGAAGAACCAGCACCAGGCTCCGGGCACCGTCCGGGCGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGGCCATTGT
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCCTCAATGCCCCACAGCCAGCACCAGCGGGCAGAGATCCACTTCTCTGAAGCAAGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCCAACC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGTGGGAGGAGTGGGAGTGGGCTGAGCTGAT
GGTCAACCAACCTGCCATCGGACCCAGCCAGCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTGCTGGT
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCTGCGGAGAAGCCACTCTGCCTGTCCA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTGCCAGGTCACTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTTAGACCACGGAACCTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCAGCAACTCTTCTTCTCTGGTGTGGTGGGGGCGAGAGAGCCATGCACTGTGAGCTGAGCGGGATGTGGGCAAGGT
CAAGTATGAGGTACGGTTTTCAACCAAGGCCAGTGCCTCAGAACCCCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGG
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGGAGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGC
TGAGAAGAAGAAAAACATCACCCCTGGACTGCGCCCGGGGCACGGCCAACCTGTGTGGTGTTCAGCTGCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTATCCAGTGTGTTACTTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCTGGTGGGT
CATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGCTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACCTGGGGCAGCCCCGCGGGAGGGCCCGGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCCTAGGTTCC
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGTGGTGTGCGCATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAAC
TCCTCCCACCCAACTTCCCCCTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT
GTAACAGGACCCCAAGGACCTGCCCTCCCCGAAGTGCTTAAACCTAGAGGCTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGTCTCTCTCTAGTTTCCCCCTCTCATCTGACCTTAGTTTGGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTATTAAAAATATTGAGAACAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMOKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFC
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLNQGDFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFYSLSGSLDMDGNQYPDLLVGS LADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPOSTQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSPDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRLDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIIETTELEVLELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNOGQSLRTLGS AFLNIM
WPHEIANGKWLLYPMQVELEGGQGPQGKGLCSRPNIHLHDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMRLDASTVIPVMVYLDPMVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPILAADGHPGLGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-
407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408
and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAACAAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACCTTGGAGTCCAGGAAAACCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATAATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAATTTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAACTTTGT
AAATCTTCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAATAAATAAATTTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKTKTLLRPERSRFDTSILPI
CKDSLGMFMNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKLLGQYIPLCEDGYYKPTQCHGSVGCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
 CTCTTCAGCCCCGGGATCGCCCCAGCAGGGATGGGCGACAAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTACACCTTCCCTCGATAGCGACTTCACCTTTACCTT
 CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCCTCGCTGGAGATCGAGTACCAAGTTTTA
 GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCCAAAACCTTAGTTTTTGAACAAAGAAAA
 TCAGATGGAGTTCACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTGAGCACCATT
 TCTGAGAAGGTGATTTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAAGAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAACTGGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAAGCAAAAGTGGGCACATACAAATCTGCTTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
 AACTTTGATAGAGTCAATTTCTGGTCTATGTTTAAATTTAGTGGTTCATGGTGGTGGTGCAGCCATTCAAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACTTAAACTCCAAACTAGAGTACGTAACATTGAAA
 AATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCATTAATGGTCTTCTCCAAAATATTTTGAGATATA
 AAAGTAGGAAACAGGTATAATTTTAAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
 TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
 TTTCTTAACCTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
 AGTCTGTTTTTAAACAGGTTCTATTACCCAGAACTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
 TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCCAA
 CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAATAAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
 AGATTTTTATAACCAATACATTTTCACTGTAACTATATAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
 CCAAAGCTGACATTTTACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
 AAATGAAGAATATAGTTTAAAGCTTCTCCTCCATAGGGACACATTTTCTTAACCTTAACTAAAGTGTAGGA
 TTTTAAATTAATGTGAGGTAAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAAATGATTGCTGACTTGGATAAATCATTATTACCAGCAGTTATGAAGGAAATA
 TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
 AGAATTTAGAAAACTTGAGAAAACCTAATCCAAAATAAAATTAAGTTAGTAAGTATAAATAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAATTTGGCTGTAGGTTTATTTTCTACAAGAATTCTGGTTTGAATTATTTTGTAAAGCAGGTACATTTTATA
 AAATGTAAGCCCTACTGTAAAGTTTACGACTGGGTGATACATATTTATTAATAAATTTTATTATAAAGTTTAT
 TAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAAACATAGACTCCCAAGTTTAA
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
 TCAAGTACTAGTAATTTAAGTTTCACTCATGAATGAAGTATAATTTTAAAGTTATGCCATTTATAACGTTGTTTAT
 GACTACATTTGTGAGTTAGAAACAACTTAAATTTGGGGTATAGAACCCTCAACAGGTTAGTAATGCTGGAATT
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTTATTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
 TCTTAGGCCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
 ATGATCAATTACCTTAATTTGGCCAAGAAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
 AGGTCAATAAGATCCTTGCCCTATGAATACCCCTCCCTTTTGGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAACTGTGATTAAGAATCTTA
 CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCT
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTAGTAAGTTCTTACCATATAAAAACGATAATTTGCTT
 TATTTGGAAGAAGATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
 CATAACCAAAAAAGCAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
 CATATGCTTTTTTTTAAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATTTAGTAAGCTGTTTATCACTT
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAAAATACTACAATATTTAAGTAAAGGCCCAACCGATTTC
 CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
 GTTATGCTGTTCTTTTCACTGTGAATGTCAAGACATGGAGGGTGTGTAATTTTATGGTAAAATTAATCCTTCTTA
 CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATTTGATGTCTCCTCAAATGAAGATTCTTTAT
 GTGAAATTTTAAAGACATTGATTCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTG
 CTCAAACCTGCTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
 AAATTTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLLAALPPVLLPGAAGFTPSLSDSDFTFLLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL
ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDVRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCTTTTTCTTACCTTCATTTAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

[illegible]

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCGGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTCTGGC
TGCTTGCAGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCACAGGAGTGT
CGGACTGTGTCAACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCCA
GCCCCGAATGCCTTGAAGAAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

protein = 43024001

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRR'TTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTTCCCTGCAGTCAGAAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACCTGATGAGTACTGCGCTAGT
CCACCCGCGGAGGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTACCAATCCCCTGCGCCGCGG
 CTGGGCCGTGCGAGAGTGCCTGTGCTTCTCTCTGACCGCGTGTGGCTCGGCCAGGCGGGTCCGCCGCCA
 GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATATTTTGTGGAATGAAAAGGA
 AGTATTAGAAATGAGCTGAAGACCATTACAGATTAAATATTTTGGGGACAGATTGTGTATGCTTGATTACCCCT
 TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTTAC
 TTAAATCAGAACTTGCAATAAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
 GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTACAGTTTACTTGG
 AGTGTCCAAAATGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
 AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
 TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
 CTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
 TGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTACTCCCCAGGCTGTTTCACTGCCATGATTTAGCTCC
 CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
 GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAATA
 TCATGGAGACAGATCAAAGGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAATTTG
 GACAGGAAATTTTGTCAACTCCATACAACTGCTTTTGTCTGGTATTGGCTGGCTGATCACTTTTTTGTTCAAA
 AGGAGGAGATTGTTTGACTTCACAGACAGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
 AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAAGTACTTTCCGCAACACACTAGAGGATCGTTT
 GGCTCATCATCGGTGGCTGTTATTTTTTTCATTTTGGAAAAATGAAAAATCAAATGATCCTGAGCTGAAAAAAT
 AAAAATCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTACTGTTCTCTGCACCAGACATCTGTAGTAA
 TCTGTATGTTTTTTCAGCCGCTCTCTAGCAGTATTTAAAGGACAAGGAACCAAGAATATGAAATTCATCATGGAAA
 GAAGATTCTATATGATATATCTTGCCTTTGCCCCAAAGAAAGTGTGAATTTCTCATGTTACCACGCTTGGACCTCAAAA
 TTTTCTGCCAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
 ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTCATGA
 GGGACTCTGTAACATGTATAACATTCCAGGCTTATCCAAACAGTGGTATTCAACCAGTCCAACATTTCATGAGTA
 TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
 ACCACCACCTTCAACGAAGTGTACACAAAGAAAAACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
 GTGTATCCTTGGCAAGTCTTAATGCCAGAAATGGAAAAAGTGGCCCGGACATTAAGTGGACTGATCAACGTTGGG
 CAGTATAGATTGCCAACAGTATCATTTCTTTTGTGCCAGGAAAAAGTTCAAAGATACCCTGAGATAAGATTTTT
 TCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATCCCTGAGAAT
 CTGGGGTCTAGGATTTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTTCAAGTGAAGAAAGTTCTACAAGG
 GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGGCAGAATTTTGTCTCCAGAATTTGAGCT
 CTTGGCTAGGATGATTAAAGGAAAAGTGAAGCTGGAAGAGTAGACTGTGAGGCTTATGCTCAGACATGCCAGAA
 AGCTGGGATCAGGGCTTATCCAAGTGTAAAGTTTTATTTCTACGAAAGAGCAAGAGAGAAATTTTCAAGAAGAGCA
 GATAAATACCAGAGATGCAAAAGCAATCGCTGCCCTTAATAAGTGAAAAATTTGAAACTCTCCGAAATCAAGGC
 GAGGAATAAGGATGAACTTTGAATAATGTTGAAGATGAAGAAAAAGTTTAAAGAAATTTCTGACAGATGACATCAG
 AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
 GAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTTATAAATATTTTA
 GACTTTGCGAGCTATAATATATGGTTACACATGAGAACAAAGATAGAGTCATCATGTATTCTTTGTTATTTGCT
 TTTAACACCTTTAAAAAATATTTAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCACTCCATG
 GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
 ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAAC
 TAGTTTTTGGTCACTTGTCTCTCTAAATAATGCTATCCCTAACCATATATTTATATTTTCTTTTTTAAACACCCAT
 GATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCTTCTCA
 AAGGTTGAAAAAATGCTTTTAAATTTTTCACAGCCGAGAAACAGTGACAGCATATATGTGCACACAGTAAGTACAC
 AAATTTGAGCAACAGTAAGTGACACAAATTTCTGTAGTTTGTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
 TAGCAATTAAGTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
 TGTGTTTATGATTTTCTGAAATGCTTTTCATAGAAATTTTCCACTGATAGTTGATTTTTGAGGCATCTAATAT
 TTACATATTTGCCCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTTCATAGTTTTG
 TTTTTCACCTCTGTCCAGTCTATTTATTTTCAAATAGGAAAAATTTTACAGGTTGTTTTTACTGTAGCTTAT
 AATGATATCTGTAGTTATTTCCAGTTACTAGTTTACTGTGACAGGGCTGCCTTTTTTCAGATAAATATTGACATAATA
 ACTGAAGTTATTTTATAAGAAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTGA
 CTCAAAGAAATCACAAATTTGTGAGTAACATGTAGTTGTTTAGTTATAAATTCAGAGTGTACAGAATGGTAAAAAT
 CCAATCAGTCAAAGAGGTCATGAATTAAGAGGCTTGAACCTTTTTTCAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDCLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLLEDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAGAIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIIHHGKKILYDILAFAKESVNSHVTTLGPPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEYEGHHS
AEQILEFIEDLMNPSVVS LTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPTQTFSEKVLQGNHWVIDFYAPWCGPCQNFAPFELLARMIKGVKAGKVDC
QAYAQCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTTCATTCCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTACCA
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTTCTTCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATTGAAATAAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTTCGAATCAATGCTGCAAAGCTTTATTTTACATTTTTT
TCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTGTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCTT
TCATGCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTGAGCATTGAAAGATTTCCCTAGCCTCTTCTTTTT
TCATTAGCCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAGTGGACCCTCTATATTTCTCCTCCCTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCCGGGGG
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGACACGGGGCCGCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCCTTCGCGAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCCGCTCGGCGC
GCGAGGCCTTCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSF SRPHTGPPPLPTPGPDRDRELT
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMI VLS
GSL LHRGAPYRDPLRIPREHVHNASAHLT FNKFWRRYGKLSRHLMKVKLK KYTKFLFVRDPF
VRLISAFRSKFEL ENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDA AQLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLF GYPK PENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

4306001

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCTGCGGCCGATGGGCCAGCCGCTGAGGTCGGGGCTGCGGCAGTGTTCTGGCCTCCGA
AGCCAACCTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCGAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGACGCCCCCGATATCCCTTCCTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGAAGTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCCTGGCCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACCGGATGAAACCGTATGCCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRI PVDLPEARCLCL
GCVNPFTMQEDRSMVSPVFSQVPVRRRLCPPPPRTGPCRQRAVMETI AVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGGCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTGAGAATGGGAGGCTTTTGC AAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACTCTCCCAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGCTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACATTTTCACAGTGACTCTTGGAATT
CCTGCTTGGTGTTCTTATGTGTTTTTCGTATAGCCACCTTGGTTTTTGGCCTTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGCCTTG TAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTTAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTTGTAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTTGAAGAAAAGTTGAATACTGTTTTGTTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCTTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQT DSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSEQRNRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

[illegible]

第 11 卷

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGS CAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLS SPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTCCGAGGAAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCTTGTGAGTCTGCTCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG
GAAGACTTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCTTTCCCTGCAGCCCACCCTGCTTCCCGA
TCCTTCCCTGACCCAGGGGCCCTTACCCTTCTGCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGTGAGTGACAAAGCCTTAGCCTCCTGCTTCCAGCACACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCGCTGTTAGCCACTTCTGTACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCCAGTGCC
GCCAGCTTACCTTCTCCTTCCACAGTCCCTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAACTGACCTCTGTGAGATTATGGGGGACATGGTGTCTTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCAGGAGCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCACA
TTGAGCAGCCCGGGCATTGGAGCAGTGTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCTGCTTCTGC
AACCCTTGACCTACTTTGCACTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTCGTCTCTGCCCTGGCCTGCCTTGTCAACATTGCCGCCTACCTCTGCTCCAGG
GTGCCCTGCCGTGCAGGAGGAACTCGGGACTACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTCT
CTGCTGGACACGAGCTTCCCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGC
ATCTTCTGCACTTCTCCCTGCTCACCTGCCTTTCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCTGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCACCTGGGCCCTCTTCAGC
CTGGTGTCTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAA
AAGTGGTGCATGTGCTGACACTGCTGGGCCTCAGCCTGGTCTTGGCCTGCCCTGGGGCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTGCTCTACCTTTTTCAGCATCATCACCTCCTTCCAAGGCTTCTCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCGGGGTGGCCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATCTAGGCCTCCAGCCACCTGCCCATGTGATGAAG
CAGAGATGCGGCCCTCGTCCGACACTGCCTGTGGCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCCAGCCGAGACT
TTGGAAGGCCAACGACCATGGAGAGATGGGCCGTGGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGTCTGGG
GGCCAGGCCTTGGATCTTGGGGTCTGGCACATCCTTAATCCTGTGCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACCTGAGGGCACTCTGCATCCTCTGTCATTTTAACCTCAGGTGGCACCACAGG
CGAATGGGGCCAGGGCAGACCTTCAAGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCTTACCTCTGAGCCCAGGCCCCCTCCCTCCCTCAGCCCCCAGTCTCCTCCATCTTCCCTGGGGTTT
TCCTCCTCTCCAGGGCTCCTTGCTCCTTCGTTACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTTGGGGA
GTGGTTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG
CTCACCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCCTA
GAGCCTGACACTCTCTAAGAGGTTCTCTCCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCCATCATGGT
TAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACCTCCTCTGCCAACATTAGTCTGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACA
GGGAGCCATCATCTCTGCTGGGAATCCTGGAAGACTTCTGCAAGGAGTCAAGCTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTACGTACCAATTCTTTTGTCTTTTGTATATTAAGAAAGTACATGTTCAATTGTAGAGA
ATTTGGAACTGTAGAAGAGAATCAAGAAGAAAAATAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAAA
AA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRFCSSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFSTFSFHSPHTAAHNASVDMCELKRDQLQLL
SQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTPVVLTQFQHLQPKNVTLCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVVSALACLVTIAAYLCSRVPPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLGLSLVLG
LPWALIFFSFASGTFQLVVLVYLFSLIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

[illegible]

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATACCAACCTGGGCCTCTTCAGCCT
GGTGTCTTCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACCTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAACTACCCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCACAGTGCAGCCTTCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCGGTGAGTGAGCTCTCACCCAGTCAGCCAAATGAGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTTACGTTTGTATGAAAGATTTGGGCCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCTATGCCACAATTACAGAAGCTGTGAGTCTTTCAGTGCTACCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGAAGACCTTATTGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACCTTCT
CAGTGTCCATAAGGGAAGAACATAAGAGAACCGATAACATTTTCTGGCCAGGTTGTCTCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCTTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTACCCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATAGAGAACGTATGCGTTATCTCCATCTTAATCTCAGTTGTTTGTCT
TCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCATAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCACCTAGCTGGGTCTGTATTTTCAGTTCTTTT
GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCCTTGCTTAAC
TCTAAAGCTCCATGTCTGGGCCATAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGTCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACATATGTTGCTATGAATTAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCATTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATTTCTCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTATGACAACCTTAGATCAACTATTTTGTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCTGTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTGAGATCCAGCCATTACTAACCTATTCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAGACTTGGCAGCTTCTGATAAAGCGTGTGCTGTGCTGTGCTGAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAGAAAATCAGTAAAATATTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATGAATCATCAAATAAAGA
ATGTGGCTATTTTGGGGAGAAAATAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLT SALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDG TILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNL L TEEVRLY
SCTPRNFVSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVT KKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
AGACTAAAAATGGTGTCTTCCAATGTGGACACTGAAGAGACAAATCTTATCCTTTTAAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCTAAAACCTCTGCCCTGTGATGTCACCTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGACAGACAAGCATTGACAGAAATCTCTGGAGGTATTTCCACGAACACCACGAACCTCACCCCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCTTTTACAGACTGGACCATCTGGTAGAGATCGATTTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTAGTGGACTCACTTATTTAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATAACCGCAGGGCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAAACATCTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTATGTTTCAATTTCAATAGAGAAA
GATGCCTTCTCTAAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATCTTGACCTAAGTGGAAATTGCCCTCGTGTATATAATGCCCCATTTCTTGTGCGCCG
TGTAATAATAATCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGGTTTTACGTCTA
CACAGTAACCTCTCTTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC
CAAACTTCTTGGCCAAAGAAATGGGGATGCTAAATTTCTGCATTTTCTCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCTTGAACAATTACATTATTTTCAATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTGTCAAGTCTCTGATTTTTCAGCATCTTTCTTCTCAAATGCCTG
AATCTGTGAGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATCCAACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCACAGCATTTGAAGAGCTTCACAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAACTTTACCAAGAACCTAAAGGTT
CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAATTAGAGGAATTAGACATCTCTAAAAATCCCTAAGTTTCTTGCCCTCTGGAGTTTTTGTATGGT
ATGCCTCCAAATCTAAAGAATCTCTCTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
CTAAAGAACCTGGAACTTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC
AGAAGCCTCAAGAATCTGATCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTCTCATGGTGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAGAGAGAAA
CATTTTAAATTTATGTTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTTGAAAACCTTTCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTGTGATGACAGACAAGTATGCAAGACTGAAAATTTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCCTCCAGCTCCGGAAGGGCTCTGTGGGAGTTCTGTCTTGTAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAGGAGAGGCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDPKKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYR
NPCYVSYSIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCNSARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMMNDNDISSSTSRTESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETDLDSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHNRFLCTCDAVWFVWVWNHTEVTIP
YLATDVTCVGPGAHKQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHYPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAAACAGAAACATGGAAAAACATGTTCCCTTC
 AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCCCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
 GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGTCAGAGTGCAGCAATCGTTCGACTACAGGAAG
 TCCCCCAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTTCATCACACACATAACGAATGAAT
 CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAATC
 CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCCTCAACCTAAAAAACCTAAGGGAGTTACTGC
 TTGAAGACAACAGTTACCCCAAATACCTCTGGTTTGGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAACA
 ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAAGTCT
 ATTTTAACAAAGTTTGCAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
 CACTATCTTTCAATTTCTTTTACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAACA
 CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAACT
 GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTCTTCAATTAATATAGATCGTTTTG
 CTTTTCAAAACCTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
 TTAATAATATGCTCATCTGAAGGTGCTGGATCTTGAATCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
 TTTTAACGATGCTGCCCCGCTTAGAAATACTTGACTTGTCTTTAACTATATAAAGGGGAGTTATCCACAGCATA
 TTAATATTTTCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCAATTTAAGAGGTTATGTGTTCCAGGAAC
 TCAGAGAAGATGATTTCCAGCCCCTGATGCAGCTTCCAACTTATCGACTATCAACTTGGGTATTAATTTTATTA
 AGCAAATCGATTTCAAACCTTTTCAAATTTCTCCAATCTGGAAATTTATTTACTTGTGAGAAAACAGAAATATCAC
 CGTTGGTAAAAGATACCCGGCAGAGTTATGCAATAGTTCCCTCTTTTCAACGTCATATCCGGAAACGACGCTCAA
 CAGATTTTGAGTTTGACCCACATTCGAACCTTTATCATTTTACCCTGCTCTTAATAAAGCCACAATGTGCTGCTT
 AGGAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTGAAAATCTTCTGACA
 TTGCTGTTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTAAAGTGAAGTGAATTTTCAGCCATTCCTCATG
 TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGTCTTACTGAATTGTCCGACTTGG
 AAGTTCTAGATCTCAGCTATAATTCACACTATTTTCAAGATAGCAGGCGTAACACATCATCTAGAATTTATTCAA
 ATTTTCAAAATCTAAAAGTTTTTAACTTGAAGCCACAACAACATTTTATCTTTAACAGATAAGTATAACCTGGAAA
 GCAAGTCCCTGGTGAATAGTTTTTCAAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
 TCTCCATTTTCAAAGGTCTCAAGAATCTGACACGCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATG
 AAGCATTCCTTAATTTGCCAGCGAGTCTCAGTGAACCTACATATAAATGATAAATATGTTAAAGTTTTTTAACTGGA
 CATTACTCCAGCAGTTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAACTACTCTTTTTTAACTGATAGCC
 TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCACCTACCTCTGGCTTTT
 TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
 AAATAAGACCACCACCAATTTATCTATGTTGGAATACACCGAAACCCCTTTGAATGCACCTGTGACATTGGAG
 ATTTCCGAAGATGGATGGATGAACATCTGAATGCAAAATTTCCAGACTGGTAGATGTCAATTTGTGCCAGTCCCTG
 GGGATCAAAGAGGGGAAGAGTATTGTGAGTCTGGAGCTTAACAACCTTGTGTTTCAGATGTCACTGCAGTGATATTAT
 TTTTCTTTCAGTTCTTTATCACACCATGTTTATGTTGGCTGCCCTGGCTCACCATTTGTTTTTACTGGGATGTTT
 GGTATATATATAATGTGTGTTTAGCTAAGGTAAAGGCTACAGGTCTCTTTCCACATCCCAAACCTTTCTATGATG
 CTTACATTTCTTATGACACCAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG
 AGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
 TCATGCAGAGCATCAACCAAGCAAGAAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACCTTTA
 AAACAGCTTTTTTACTTGGCTTTGAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCTGCTGGAGC
 CAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA
 ACCCGAAGGCAGAAGGCTTGTGTTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAATGATTACGGGTATAACA
 ATATGTATGTGATTCATTAAAGCAATACTAACTGACGTTAAGTCATGATTTTCGCGCCATAATAAAGATGCAAG
 GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAAACCTTAGTGGTTTTAAACAACACA
 TTTGCTGGCCACAGTTTTTGGAGGTCAGGAGTCCAGGCCAGCATAACTGGGTCTCTGCTCAGGGTGTCTCAG
 AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
 ATTCCTCCTGGGCTATTGGCCAAAGGCTATACCTCATGTAAGCCATGCGAGCCTCTCCACAAGGCAGCTTGCTTC
 ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAAATCGAATCAAAAAAGTGAT
 ATCTCATCACTTTGGCCATATTCTATTTGTTAGAAAGTAAACCACAGGTCCACCAGCTCCATGGGAGTGACCACC
 TCAGTCCAGGGAAAAACAGCTGAAGACCAAGATGGTGAGCTCTGATTGCTTCAGTTGGTCATCAACTATTTTCCCT
 TGACTGCTGTCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
 ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGCCACTGTGACTAATGGTCTTAATA
 TTAAGCTGTTGTTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTTGGTTTTAT
 TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
 ATAGCTTTTAAAGCATCTTTTACTTCTTACCAATTTTAAAGATATGCAGCTAAATTCGAAGCTTTTGGTCTATA
 TTGTTAATTGCCATTGCTGTAAATCTTAAATGAATGAATAAAATGTTTTCAATTTTACAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQND SVIAECSNRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFQNLTLQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNLYVGEIVSGAFLTMLPRLEILDLSFNIIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV
LDLSYN SHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTE LHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKI PRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWD PGLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTA FYLALQRLMDENMDV IIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCCTG
GAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTGA
AAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGGCCCCGGGGAGGGGAAGTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCAGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGACGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCTGGGGTCTTGCTCCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCTCCTCCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKDS

Signal sequence:

1-19

bioRxiv preprint doi: <https://doi.org/10.1101/100000>; this version posted October 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCCTGGAGGAGCAGCT
GGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAAATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSDGTLCVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

bioRxiv preprint doi: <https://doi.org/10.1101/100000>; this version posted April 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 218

GGTTGCCACAGCTGGTTTATGGGCCCCGACCCTGGGGCCCCCTTGTACAGGAGGAGACAGCCTCCCGGCCCGGGGAG
 GACAAGTCGTGCCACCTTTGGCTGCCGACGTGATCCCTGGGACGGTCCGTTTCTGCGGTACAGCTGCCGGGCCG
 AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGGTCTCCCTTCTCCGCTGGGCGGGTTATCGGGAGG
 AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
 TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCAGGCAGGAACACCTTTTGCTGTG
 ATGGCCCGCTCATGATGGCCCCGGCAAAAGGGCATTCTTCTACCTGACCCTTTTCTCATCCTGGGGACATGTACAC
 TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTTCTGCTGTCTCTGCCATCCCTGTATTTGCTGCCATGCTCT
 TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCACTGACCCTGGAGTGATTCTCGGGCGCTACCAGATG
 AAGCAGCTTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGGGTGCCCCAGGGCCAGCGACACCGCCTCGTA
 TCAAGAATTTCCAGATAAACAACCAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
 GGGCCTCCCATTTGCAGCATCTGTGACAACGTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTG
 TTGGAAAGAGGAACCTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCTCACATCTATGTCTTCGCCT
 TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAATTTGGCTTCTTGAGACATTGAAAGAAACTCCTGGAACCTG
 TTCTAGAAGTCTCATTGCTTCTTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTTATACTTTCTCTGTGG
 CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
 ATGGCAATATTGTGAAGAACTGCTGTGAAGTGTGTGTGGCCCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA
 TTTTGCCACTGGAGGAAAGTGAAGTGCACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
 CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
 AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCCTATCTATGGAAGAGACTTTTGTGTTGTGTT
 TAATTAGGGCTATGAGAGATTTAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
 GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTCTTTCTGTGCAAGCTTTTTTAAATTTCTGAACT
 CAAGGCAGTGGCAGAAGATGTGAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
 CCATGGCCTCAGCCACAGGGTCCCCCTTGACCCCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTCAC
 TGGTCTCATTCTGGGGCTAAAAGTTTTTGGAGCTGGCTCAAATCTTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
 GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
 TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCCACAAATCCTTTTAGGAATGGGACAGGTACCT
 TCCACTTGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTGGACTCCTGCTCCCATTAGGAG
 CAGGAATGGCAGTAATAAAAGTCTGCACCTTTGGTCACTTTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
 ACTATCCCCTCAGACTCCCTGTGTGAGGCCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
 GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAAGCCT
 CGGCTGAGTGAGGGAAAGCCAGCACTGTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCT
 ATGGTAACCACACTGGGGGCTTCCCTCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTT
 CACCCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTAGGGAAGAAGATTTATGT
 ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
 GCGGTGGGGGAGTGTAACCGGAACCTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFC CDGRVMMARQKGI FYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLCFFTLWSVVGLTGFTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCGTGAACTATCCACCATA
TTTCAGAAGCCAAGGGTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNVVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYECsasNDVAAPVVRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNITCVASNKLGHTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGCCGCGGCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCGACTCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACA
GAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGCTGGATTTCGGAGGAACCCCTGCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTTACCCTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCACTATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGAG
CAGGTGTGAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACCTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACCTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATC
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG
GCTGCCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTTGCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCCGCTTCCGCGTGCTGGGCACCTCTCTCCAACCTCCCGTGACTTCCTGCGG
CACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTGAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGTCTCTTGGGTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG
TGCCCTCTGCTTTGGGGGTGCCCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTACCCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESIDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFFTVYISADSKSSNSNVIQVDQSGFLFLPSRDYYLNRTANEKVLTA
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLLETLYGTTKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPVAVGLTNHQLFFVVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 226

GCCCCGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCGCTCCCGCGCCCTCCTCCCTCCCTCCTCCC
CAGCTGTCCCGTTTCGCGTCATGCCGAGCCTCCCGGCCCGCGCGGCCCGCTGCTGCTCCTCGGGCTGCTGCTGCT
CGGCTCCCGGCCCGGCCCGCGCGGCCCGGCCAGAGCCCCCGTGTGCCATCCGTTCTGAGAGAAGGACCGCTGCC
CGTTCCGGGAGCGGCAGGTAGGTGGGCGCCCGGGGAGGCGCGGGCGGGGAGTCCGGGCTCGGGGCGAGTCAGCGC
CAGCCCCGAGGGGGCGCGGGGCGAGGTGGCTCGGCGCGCGGGCGGCCCGAGGGTGGGCGGGGGCAGAAGGGC
GCGGTGCTGGGACCCCGGACCCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGCTGGGCGAGCGGCCTCCAGC
CAAGCCCGTCCCGCAGGCTGCACCTTCGGCGGGAAGGTCATGCCCTTGACGAGACGTGGCACCCCGACCTAGG
GGAGCCATTCCGGGTGATGCGCTGCGTGTGTGCGCTGCGAGGCGCAGTGGGGTCCCGTACCAGGGGGCCTGG
CAGGGTCAGCTGCAAGAACATCAAACAGAGTGTCCCAACCCCGGCTGTGGGCGAGCCGCGCCAGCTGCCGGGACA
CTGCTGCCAGACCTGCCCCAGGACTTCGTGGCGCTGCTGACAGGGCGGAGGTGCGAGGCGGTGGCACGAGCCCG
AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTT
CTCAGACTCCAATGGCAGTGTCTGTTTGAGCACCTGCGAGCCCCACCAAGATGGCCTGGTCTGTGGGGTGTG
GCGGGCAGTGCCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCA
CCCTTCAGGGGAGGTCTGGGGGCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGAC
TCTAGAAGGCCCCCAACAGCAGGGCGTAGGGGGCATCACCTGCTCACTCTCAGTGACACAGAGGACTCCTTGCA
TTTTTTGCTGCTCTTCCGAGGCCTTGACAGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACAGGGGCA
GCTACTGCGAGAATTCAGGCCAATGTCTCAGCCCAGGAACAGGCTTTGCTGAGGTGCTGCCCAACCTGACAGT
CCAGGAGATGGACTGGCTGGTGTGGGGGAGCTGCAGATGGCCCTGGAGTGGGCGAGGCGAGGCCAGGGCTGCGCAT
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGTCCTGCAAAGTGTCTTTGTGGGGCTAATGCCCTGATCCC
AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCTGATCCTCCAGGTGCAATT
GGTAGGGACAACCAAGTGAAGGTGGTGGCCATGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCCCACTGTCT
GTGCCACATGGCTGGCCTATCCTCCCCGAGGCCGTGGGTATCTGCCCTGGGCTGGGGTGGCCGAGGGGCA
TCATATGCTGCTGCAGAATGAGCTCTTCTGAACGTGGGCACCAAGGACTTCCAGACGGAGAGCTTCGGGGGCA
ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGGCCCTAGCAGGAGCCCTGGTGTACC
CCCTGTGAAGAGCCAAGCAGCAGGGCACGCCTGGCTTCTCTTGATACCACTGTCACTGCACTATGAAGTGCT
GCTGGCTGGGCTGGTGGCTCAGAACAAAGGCACTGTCACTGCCACCTCCTTGGGCTCCTGGAACGCCAGGGCC
TCGGCGGCTGCTGAAGGATTCTATGGCTCAGAGGCCACGGGTGTGGTGAAGGACCTGGAGCCGGAAGTGTGCG
GCACCTGGCAAAGGCATGGCTTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCT
CTCCTCCCAGGTGCACATAGCCAACCAATGTGAGGTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAGGGGGT
GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCGCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCGC
CAAACCTGGTGGTCTGGGCGGCCCGGAGACCCCAACACATGCTTCTTCGAGGGGCGAGCAGCGCCCCACGGGGC
TCGTGGGCGCCCACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGTCTGTGACCCGGT
GGTGTGCCACCGCCAGCTGCCACACCCGCTGCAGGCTCCCGACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT
TTTTGATGGTGACCGGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTGTGCTGCTGCTGCTGCTGCTGCT
GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCACTGTCCCCGGCTGGC
CTGTGCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCAACAGTGTCCAGGTGAGGCCCCACCCCGAGCTGGG
GGACCCCATGCAGGCTGATGGGCCCCGGGGCTGCCGTTTGTGCTGGGCGAGTGGTTCACAGAGAGTCAGAGCTGGCA
CCCCCTCAGTGCCCCGTTTGGAGAGATGAGCTGATACCTGCACTGAGATGTGGGGTAAGTGGGGAGCAGAGGCTGT
GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGCTTCCAGAGGAGTCCCTGAAGAAGCTGAAGGTCA
CTGTGTCCAGTGCCTCTGGGGGACACTCAGTGTCTGCTCTGTCTTGTACCAGGCAGGGGTGCCTCACTGTGAGC
GGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCACC
GGCGGCGTAAGTGAGGGAGTCCAGGGTCAGCAGCTGTGAGTGGAGGGCTCACTGCCTGTGGGACTCCTGATCAG
GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCAAAAGTGCCATTCCAATCCACCTCACAGCA
ACCTGGTGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTCTGAAGCTCAGAGAAATTAAGCAACGAG
ATGAAGGTCAACAGCTGTGTGCACTGACCTGTTTAGAAAAATACTGGCCTTTCTGGGACCAAGGCAGGGATGCTT
TGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCACTCCCTCTCCCTCTCCCAACATTCCCTCCCTTCTGTCTCC
AGCAGCCCCAGAGACCAGAATGATCCAGAGCTGGAGAAAGAACCGAAGGCTCTTAGGGAGCAGCCAGAGGGCC
AAGTGACCAAGAGGATGGGGCCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCACTTCTCCTGTGGGAAG
CCCAGTGCCTTTGCTCCTCTGTCTGCTTACTCCCACCCCACTACCTCTGGGAACCAAGCTCCACAAGGGG
GAGAGGCAGCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCCCCTGCCACCTCGGCCCTCTGTCTGGAA
GCCCCACCCCTTTCTTCTGTACATAATGTCACTGGCTTGTGGGATTTTAAATTTATCTTCACTCAGCACCAAG
GGCCCCGACACTCCACTCCTGCTGCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTTGTATTTATTTAAAC
ATTTCTTTTTTCAGTCTTTGGGCATGAGGTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCCTGGTGGAGAAGGG
GCNGAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
CGTGGCANNTTGGCTGGCATNCTGGGTTCGCGAGAGGGGCTGGGGATGGTCTTGAGATGGTCTAGAGACTCAAG
AATTTAGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCACATCGCTGGCCTGTTTGTGACTTCATG
TTTGAAGTTGCTCCAGAGAGAGAATCAAAGGTGTCAACAGCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTCT
TTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCCGCGCCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTGCGA
GGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCCT
CCTTTCTTTCTTCTTTTCTTCCCTCCCTCCCTTTCTTCCCCTTTTCCCTTCCTTCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAAC TCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTGCGTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCGGTGTACCCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCCAAAAACACCGAAAAACAAAAAGAGAGAAACAACACCCAACTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACAGTGTCTGTTTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCAGGAAAATGCTGTGTCAAT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTGTTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAACTTACAACCTTTTCAAAGCATTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTGAGTTCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTCTTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTGAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAATAATATAAAATGTTTGTCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCCTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATAACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCTTTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCTTTTCACTAGTCCAAGCCAAAAA
TTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LSLCCLLP SCLPAGQSVDFPWA AVDNMMVRK GDTAVLR CYLED
GASKGAWLNRSSIIFAGGDKWSVDPRVSI STL NKR DYS LQIQNV DVTDDG PYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTLCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFPDVRKV KVVVNFAPT IQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLK NAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

Parameter	Value	Standard Error	z	P	95% CI
Intercept	1.00	0.00			
Age	0.01	0.01	0.10	0.92	-0.01, 0.03
Gender	0.01	0.01	0.10	0.92	-0.01, 0.03
Education	0.01	0.01	0.10	0.92	-0.01, 0.03
Income	0.01	0.01	0.10	0.92	-0.01, 0.03
Health status	0.01	0.01	0.10	0.92	-0.01, 0.03
Family size	0.01	0.01	0.10	0.92	-0.01, 0.03
Marital status	0.01	0.01	0.10	0.92	-0.01, 0.03
Religious beliefs	0.01	0.01	0.10	0.92	-0.01, 0.03
Community support	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare access	0.01	0.01	0.10	0.92	-0.01, 0.03
Health insurance	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare quality	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare cost	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare availability	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare accessibility	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare utilization	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare satisfaction	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare quality of care	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient safety	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient satisfaction	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient engagement	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient participation	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient empowerment	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient autonomy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-efficacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-management	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-monitoring	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-education	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-reflection	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-regulation	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-assertion	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-advocacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-empowerment	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-efficacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-management	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-monitoring	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-education	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-reflection	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-regulation	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-assertion	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-advocacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-empowerment	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-efficacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-management	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-monitoring	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-education	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-reflection	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-regulation	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-assertion	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-advocacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-empowerment	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-efficacy	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-management	0.01	0.01	0.10	0.92	-0.01, 0.03
Healthcare patient self-monitoring	0.01	0.01	0.10	0.92	-0.01, 0

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCTCTTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTTCACTCAGAACCAGGGATGTTTCAAGATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCCTCCAAGGACCCAGGGTGTCTCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAAGGGGAAACTGGAACCTAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCCTCCTGGAGCCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTCTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCGGGGCCCTCCTGGTGCAGTGGGACACCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTTCAGTGTCGGGGCACGGAGAGTACCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGAACCGGAAACCTTTTCA
CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAPHQIAMEPF EINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDGLGLPGSKGDRGMKGDAGVMGPPGAQGSKGDGFRPGPPGLAGFPGAKGDQGPGLQGVPG
PPGAVGHPPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGGLQGGQGRKGESGVPGPA
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGTCCAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCG
TGATTCCTTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAACACCATTCCCTGATTTT
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1 .

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG

KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCAGCCTGCAGGG
 CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCGTGGTGGTTGGAGGGC
 GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
 TGGAAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCGCGCTGGCTGTG
 CGCTGGGGCGCTGGTGTGCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
 TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAGCATAATATGAAAGCATTTTTGGAT
 GAATTGAAAGCTGAGAACATCAAGAAGTCTTACATAATTTTACACAGATACCACATTTAGC
 AGGAACAGAACAAAACCTTTCAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCC
 TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTGTCTACCCAAATAAGACTCATCCC
 AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCTC
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
 TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTTGTAATTGCCAGATATGGGAAAGT
 TTTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
 CCGACCTGTCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
 GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
 AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
 GTATTCCTGTTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
 TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
 CTTTACTGGAACTTTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
 CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
 CTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
 TGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
 CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA
 GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
 AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
 TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
 ACTAAAAAAGTCCTTCCCCAGAGTTCAGTGGCATGCCCAGGATAAGCAAATTGGGATCTGG
 AAATGATTTTGGAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTA
 AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
 GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTTCG
 AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTGATTGTGAGATTATG
 CTGTAGTTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
 ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
 TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
 TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGAC
 AGGCCTTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
 ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
 GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
 TTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
 CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAATTTGGTATATTTGAAATAAAGT
 TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIRYGVFRGNKVKNQAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQQRGNILNLNGAGDPLTPGYPANAYRRGIAEAVGLPSIPVHPIGYIDAQKLLKMG
GSAPPDSSWRGSLKVPYNVGPFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713